ORIGINAL RESEARCH **Open access**

Multirate Sampling-Based Energy Efficiency Analysis in Petrochemical Plants via Convolutional Neural Networks

Laura Keller[1*], Thomas Lehmann[1], Simon Brunner[2], Sophie Laurent[3]

Abstract

Optimizing energy efficiency within the petrochemical sector plays a vital role in curbing overall energy usage and mitigating ecological degradation. Accurate modeling of energy efficiency provides critical insights needed to guide operational modifications. In industrial settings, uneven data collection frequencies often compromise the reliability of conventional energy efficiency evaluation methods, making multi-rate data challenging to manage. To address this issue, this study introduces a multi-channel convolutional neural network paired with a parameter-based transfer learning framework designed to optimize energy efficiency forecasting amidst irregular sampling periods. The multi-channel configuration isolates distinct temporal patterns by applying convolution across the time domain. Simultaneously, transfer learning empowers the network to adapt to unfamiliar input distributions post-training. The model's predictive accuracy and capacity for conservation analysis are verified through a case study focused on vinyl chloride monomer manufacturing. The findings demonstrate that the developed architecture surpasses conventional baselines and standard convolutional structures regarding precision and consistency, achieving an R-squared of 0.97. Incorporating transfer learning successfully mitigates performance degradation, upgrading the model's adaptability during continuous energy monitoring. Furthermore, assessing the variance between predicted and actual energy usage reveals substantial optimization opportunities, capable of lowering annual energy demand by a mean of 7.25% and decreasing carbon dioxide emissions by 5,709 tons.

Keywords Transfer learning, Energy efficiency prediction, Petrochemical process, Multirate prediction, Convolutional neural network

*Correspondence:
Laura Keller
laura.keller@ethz.ch
[1] Department of Artificial Intelligence for Smart Grid Optimization, Faculty of Electrical Engineering, ETH Zurich, Zurich, Switzerland
[2] Department of Intelligent Energy Systems and Automation, Faculty of Engineering, EPFL Lausanne, Lausanne, Switzerland
[3] Department of Artificial Intelligence and Sustainable Energy, Faculty of Engineering, Sorbonne University, Paris, France

Introduction

Worldwide energy requirements have risen steadily from the nineteenth century onward [1], with industrial operations contributing 24.2% of global greenhouse gas outputs [2]. Reports on industrial energy use from the Energy Information Administration indicate that chemical processing represents one of the most energy-intensive sectors [3], accounting for 37.77% of total energy consumption in 2021 and yielding roughly 500 million metric tons of output annually, as illustrated in **Figure 1**, where expenditures on energy and feedstock constitute approximately two-thirds of manufacturing outlays [4]. The petrochemical sector, a key branch of chemical production, faces escalating demands to conform to sustainable development principles and attain net-zero emissions. Consequently, industry efforts have shifted toward modeling energy efficiency and identifying large-scale conservation potential to curtail expenditures and environmental strain [5].

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, distribution, reproduction, and adaptation in any medium or format, provided that appropriate credit is given to the original author(s) and source, a link to the license is provided, and any changes made are indicated. Adapted material must be distributed under the same license.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

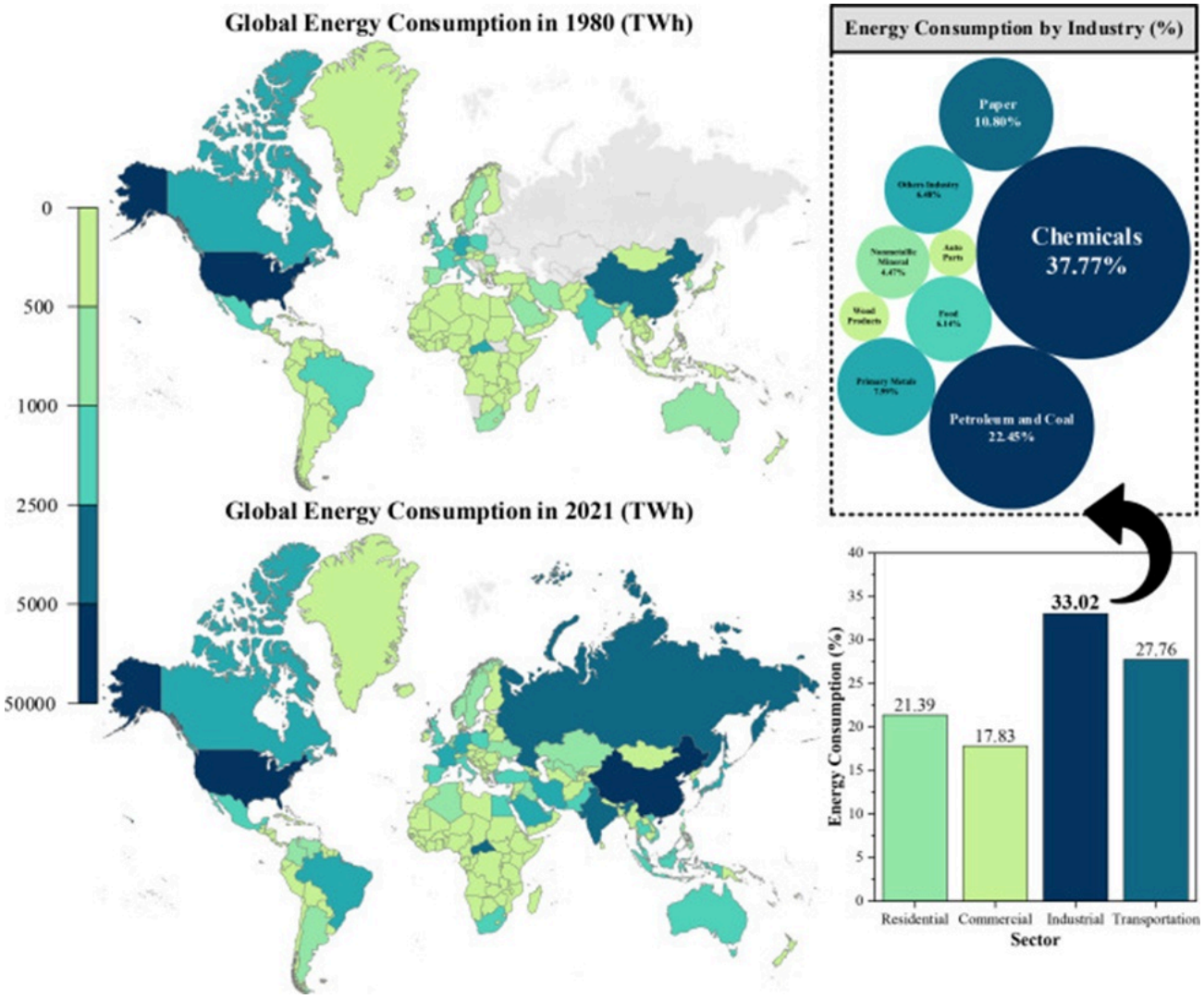


Figure 1. Energy Consumption by national, sector, and industry types.

Driven by the rapid expansion of data-centric methodologies in recent times, a variety of scholarly techniques have been implemented for energy tracking and conservation assessment. For instance, analytical mathematical formulations, which demand a precise comprehension of variable interdependencies, have rarely undergone comprehensive validation [6]. While principal component analysis (PCA) remains an option for tracking energy utilization, it struggles to establish an adequate mapping dimensionality [7]. Data envelopment analysis offers a mechanism to monitor efficiency and guide operational optimization to some degree [8], yet it experiences limitations in managing parameter variance when confronted with ambiguous datasets [9].

To navigate data complexity, multiple investigators utilize artificial neural networks (ANN) because of their versatile training algorithms, straightforward deployment, and strong performance in chemical operations [10]. For short-term demand forecasting in industrial settings, Walser and Sauer [11] deployed convolutional neural networks (CNN). Machalek *et al*. [12] developed a gated recurrent unit (GRU) using an encoder-decoder framework to execute moving-horizon energy projections for industrial boiler networks. Liu *et al*. [13] introduced an attention-driven GRU framework, where the integration of an attention layer allows the system to prioritize critical segments of the input array. Predictive capabilities can also be augmented via residual networks (RESNET); such frameworks incorporate skip connections within internal residual units to facilitate learning from both original inputs and intermediate features, though they remain vulnerable to vanishing gradient complications [14].

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Thanks to its specialized internal gating architecture, the long short-term memory (LSTM) model can successfully counter gradient issues [15]. These LSTM systems have effectively modeled the transient behavior of fluctuating disturbances [16] and shifting parameters [17]. Nevertheless, as the observation timeline lengthens, the capacity of LSTM structures to sustain long-range dependencies diminishes, thereby degrading model reliability. Transfer learning has expanded as a valuable strategy to boost the diagnostic proficiency of neural networks [18] and lower the volume of training data needed [19]. This technique not only facilitates the adjustment of pre-trained architectures to novel objectives or target domains by calibrating model weights using compact target datasets [20] but also strengthens operational adaptability when input data characteristics deviate from the initial training environment [21].

However, when analyzing broad system boundaries, networks frequently confront high-dimensional challenges stemming from plantwide variables [22]. CNNs are regularly chosen to process multi-dimensional data arrays because of their robust feature extraction qualities [23]. Convolutional kernels isolate different regions of input information, while the localized connectivity of a CNN identifies regional trends within the data and consolidates them into abstract, higher-level descriptions [24]. To isolate a broader array of features, a multi-channel CNN can achieve superior pattern identification relative to standard single-channel variants by capitalizing on the diverse inputs across separate channels [25]. However, such systems typically lack intrinsic adaptive tracking capabilities. Furthermore, artificial intelligence (AI) has been widely adopted to construct energy frameworks for diverse purposes, including optimization, policy evaluation, and predictive maintenance aimed at eliminating energy wastage [26]. Even so, given the specific criteria and forms of data needed, only a minimal number of implementations suit chemical engineering environments [27].

Due to operational safety protocols, constraints on instrument deployment, and inherent process features, collecting measurable process parameters at uniform intervals can be highly problematic, particularly in heavy industry [28]. Data collected from chemical processes are uniquely sparse, restricted, and variable [29], which undermines the effectiveness of conventional machine learning paradigms in chemical engineering applications [30]. Petrochemical operations frequently integrate qualitative and quantitative indices, necessitating offline laboratory testing to derive supplementary features or verify ongoing values. Consequently, the observed parameters display distinct characteristics and irregular sampling frequencies [31].

In light of the current literature reviewed above, most investigations conduct energy diagnostics using consistent sampling rates and non-deficient datasets, leaving the challenges of multi-rate intervals unaddressed [32]. While density-based spatial clustering of applications with noise (DBSCAN) can reconstruct deficient datasets into uniform sampling rates [33], this technique substantially diminishes the volume of accessible records, thereby inducing a sparse and inadequate data environment. As a consequence, the practical efficacy of AI-driven, plantwide energy efficiency and savings diagnostics that account for multi-time-scale data constraints has not been thoroughly examined [34], establishing the core rationale for the research detailed in this study.

Accordingly, this paper seeks to bridge the research deficit regarding the creation of real-time energy efficiency prediction systems alongside the identification of conservation opportunities under multi-time-scale data challenges by leveraging a multi-channel transfer learning convolutional neural network (TLCNN) framework. When process information cannot be gathered at identical sampling frequencies, conventional energy efficiency computations require a fully completed vector of input parameters, preventing instantaneous online evaluation. By implementing the TLCNN architecture, this investigation demonstrates that the AI system can successfully accommodate variations in process mechanics or data distributions post-training to forecast energy metrics in the petrochemical sector. Furthermore, the proposed strategy operates effectively without requiring prior data reconstruction or reducing available data records. The vinyl chloride monomer (VCM) manufacturing process serves as the industrial case study to validate the network's utility and stability. This research provides a fresh perspective on prognostic design and model engineering for investigators applying AI strategies to elevate energy monitoring performance in intricate petrochemical setups by mitigating multi-rate sampling difficulties, with the core innovations outlined below:

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

1. Present a novel technique for estimating energy efficiency in scenarios where uneven sampling periods, e.g., 1, 5 and 10 min, occur within an industrial petrochemical case study, demonstrating how a multi-channel configuration empowers individual CNN channels to isolate distinct data patterns.

2. Pinpoint the energy-critical variables that define plantwide production dynamics within the VCM process. A hybrid approach combining tree-based feature isolation and forward stepwise selection is deployed to determine highly correlated input features, preventing superfluous associations between inputs and outputs during network training.

3. Devise a TLCNN framework to upgrade adaptive training proficiency when data collection rates shift post-deployment and deviate substantially from the initial training baseline. Furthermore, the outcomes confirm that transfer learning significantly increases training consistency within a multi-time-scale predictive context.

4. Appraise the estimated energy variance to determine potential conservation savings, moving toward a more sustainable and ecologically conscious operation.

Model Development Framework

Figure 2 demonstrates the structural flow of model construction, encompassing feature selection, data segmentation, network architecture design, error assessment, and diverse evaluation scenarios. The specific components of each phase are outlined in the subsequent subsections.

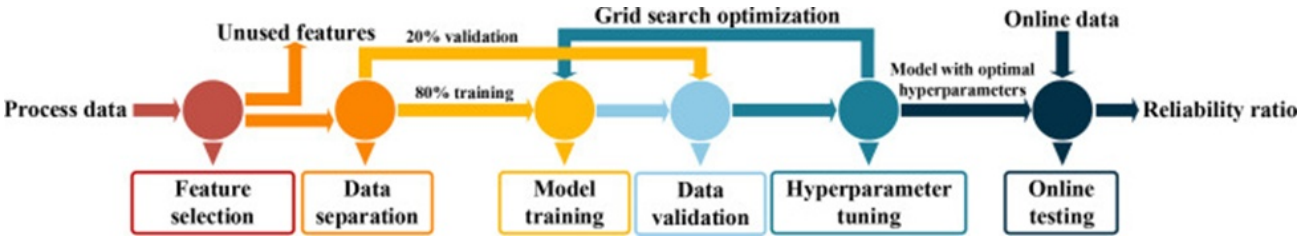


Figure 2. Model development of this study.

Feature selection

For extensive industrial operations, estimating energy performance demands sophisticated feature engineering strategies to isolate highly interdependent variables and features that offer substantial conservation opportunities [35]. The screening methodology executed in this research integrates a dual-stage framework. The initial phase leverages tree-based architectures to quantify the significance of various indicators, utilizing decision trees as an embedded screening tool well-suited for isolating critical inputs [36]. This derived metric reflects the relative weight each explanatory variable holds in projecting the target variable, which represents energy efficiency within this framework.

Calculating the feature importance (FI) involves aggregating the variation at each node caused by splits across each independent variable, subsequently normalizing this value by the cumulative sum of all tree branches. Elements failing to meet importance thresholds are discarded. The mathematical formulation defining feature importance is expressed in **Eqs. 1 and 2.**

$$NI_j = w_j e_j - (w \cdot e)_{child\,j} \quad (1)$$

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

$$FI_i = \frac{\sum_{j \in \text{Nodejsplitonfeaturei}} NI_j}{\sum_{k \in \text{allnodes}} NI_k} \quad (2)$$

Here, wj represents the sample weighting factor assigned to node j, ej denotes the node's localized error, NIj indicates the specific node significance, FIi signifies the importance metric for individual feature i, and k spans the complete set of network nodes. Essentially, a higher FI metric indicates a stronger correlation between the chosen variable and system energy efficiency.

In the subsequent phase, a forward stepwise algorithm initiates model training by selecting the highest-ranked FI attribute from the preliminary screening. Successive iterations sequentially incorporate the remaining parameters according to their descending priority score. These cycles continue until all ranked parameters are evaluated. Ultimately, the optimal subset of variables yielding the peak adjusted R^2 value is established as the primary input vector for model benchmark evaluations. The overarching feature distillation and elimination architecture is mapped in **Figure 3.**

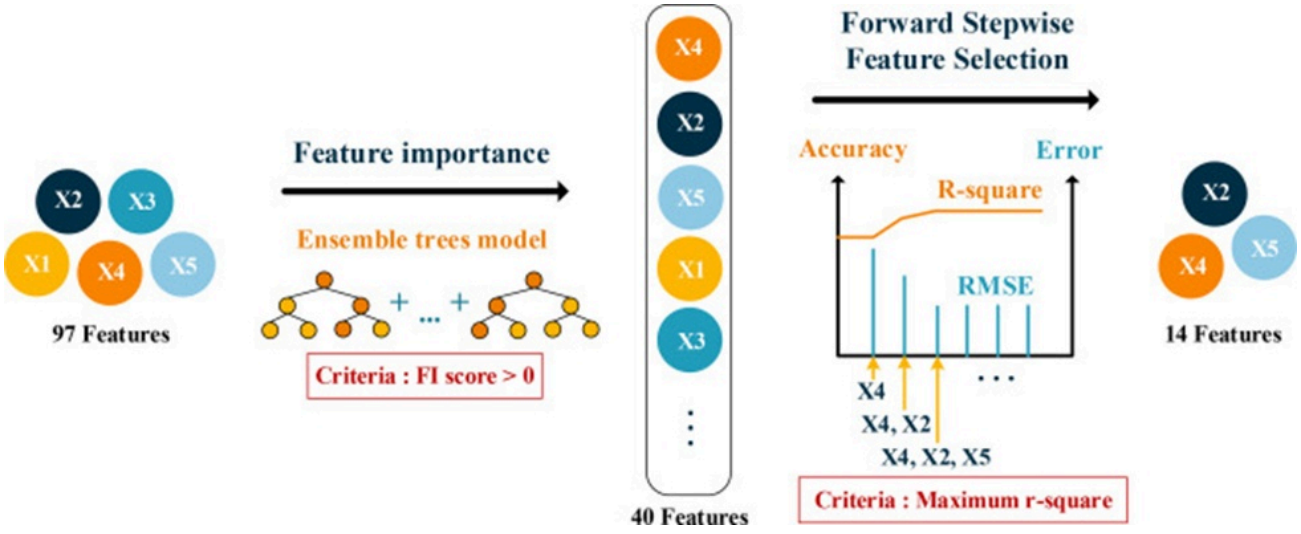


Figure 3. Framework of feature selection and elimination criteria in each step.

Multi-channel transfer learning convolutional neural network (TLCNN)

The proposed TLCNN configuration utilizes five primary functional tiers: an input gateway, a convolutional stage, a spatial reduction layer, a dense layer, and an output layer. This structural topology is diagrammed in **Figure 4**, while the internal trainable parameters are compiled in **Table 1**. The operational role of each constituent layer is detailed below.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

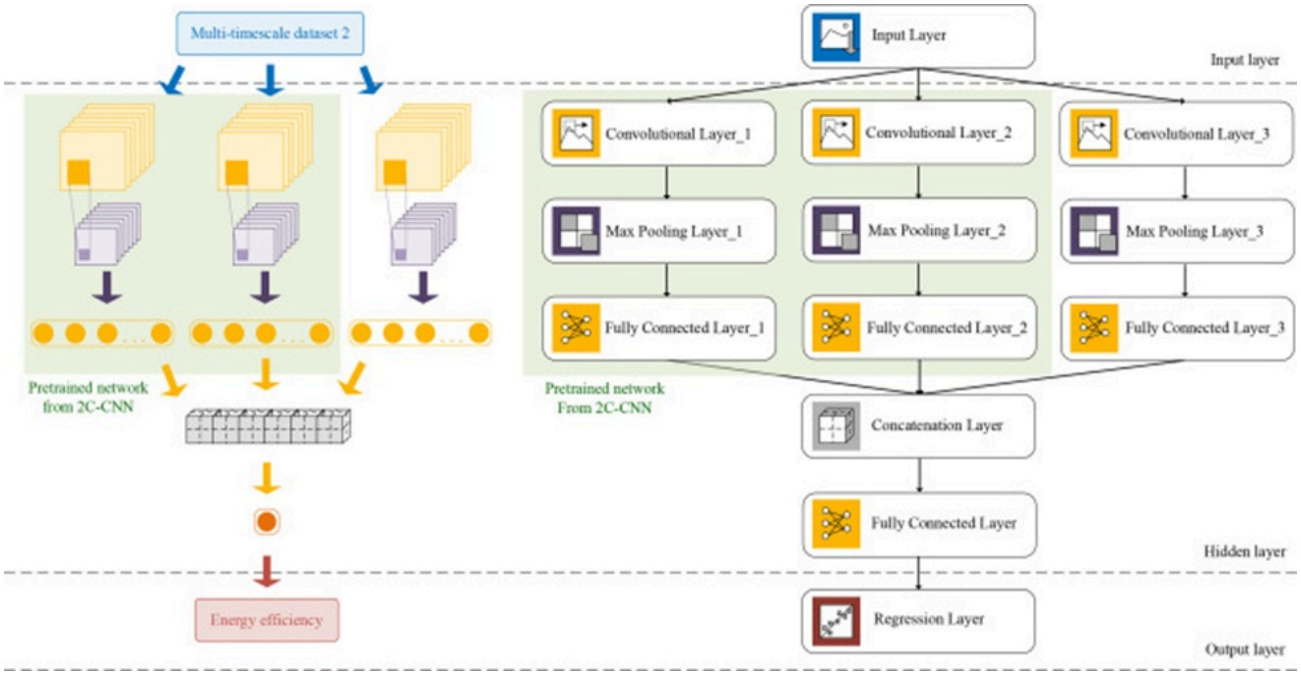


Figure 4. TLCNN model structure.

Table 1. The characteristics of the TLCNN model.

Layer name	Layer type	Activations	Learnable parameters
Input layer	Sequence input	14 × 1 × 1	–
		14 × 1 × 1 with *z-score* normalization	
Convolutional layer_1	Convolution	14 × 1 × N	Weights: 10 × 10 × 1 × N
	50 filters of size 10 × 10 × 1		Bias: 1 × 1 × N
Max pooling layer_1	Max Pooling	14 × 1 × N	–
	3 × 3 max-pooling operation		
Fully connected layer_1	Fully Connected	1 × 1 × N	Weights: N × 700
	50-neuron fully connected layer		Bias: N × 1
Convolutional layer_2	Convolution	14 × 1 × N	Weights: 10 × 10 × 1 × N
	50 filters of size 10 × 10 × 1		Bias: 1 × 1 × N
Max pooling layer_2	Max Pooling	14 × 1 × N	–
	3 × 3 max-pooling operation		

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Fully connected layer_2	Fully Connected	1 × 1 × 1	Weights: N × 700
	Single fully connected layer		Bias: N × 1
Convolutional layer_3	Convolution	14 × 1 × N	Weights: 15 × 15 × 1 × N
	50 filters of size 15 × 15 × 1		Bias: 1 × 1 × N
Max pooling layer_3	Max Pooling	14 × 1 × N	–
	5 × 5 max-pooling operation		
Fully connected layer_3	Fully Connected	1 × 1 × N	Weights: N × 700
	40-neuron fully connected layer		Bias: N × 1
Concatenation layer	Concatenation	1 × 1 × 3	–
	Combination of 3 channels along dimension *t*		
Fully connected layer	Fully Connected	1 × 1 × 1	Weights: 1 × N
	Single fully connected layer		Bias: 1 × 1
Regression layer	Regression Output	1 × 1 × 1	–

*N is the final hyperparameter.

Input layer

This initial network threshold functions without structural weights or operational biases, serving primarily to standardize incoming data streams. We employ z-score normalization at this entry point to project all features onto a uniform scale, thereby anchoring the stability of the training process via **Eqs. 3–5.**

$$z_i^j = \frac{x_i^j - \mu^j}{\sigma^j} \quad (3)$$

$$\mu^j = \frac{\sum_{i=1}^{n} x_i^j}{n} \quad (4)$$

$$\sigma^j = \sqrt{\frac{\sum_{i=1}^{n} \left(x_i^j - \mu^j\right)^2}{n-1}} \quad (5)$$

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Where x_i^j represents the i-th raw observation belonging to the j-th input profile, μ^j corresponds to the calculated mean of that specific feature, and σ^j indicates its associated standard deviation.

Convolution layer

Trainable elements within this layer extract latent patterns by executing mathematical convolutions over the input array. Broadly defined, convolution combines the raw dataset with a localized scanning filter; for unidimensional arrays, this scanning occurs along the temporal axis. Here, the primary regression framework incorporates convolutional nodes to construct an abstract feature space, enabling automated, unsupervised identification of dominant signal attributes. The resulting convolutional state is computed according to **Eq. 6.**

$$y_j = f(x_i * w_k + b_k) \quad (6)$$

Where f acts as the designated non-linear activation operator, wk represents the weight matrix, and bk defines the native bias of the scanning filter.

Pooling layer

Following feature extraction, a down-sampling layer compresses data dimensionality by consolidating clusters of active neurons into a unified node for the next processing stage. Lacking internal weights or biases, this step shrinks the informational footprint of individual feature maps, effectively reducing the volume of learnable metrics in downstream layers to accelerate convergence and thwart overfitting. Common strategies include maximum and average pooling; the former captures the peak amplitude within a spatial window, whereas the latter computes the arithmetic mean. The pipeline connecting the convolutional and down-sampling operations is illustrated in **Figure 5.**

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

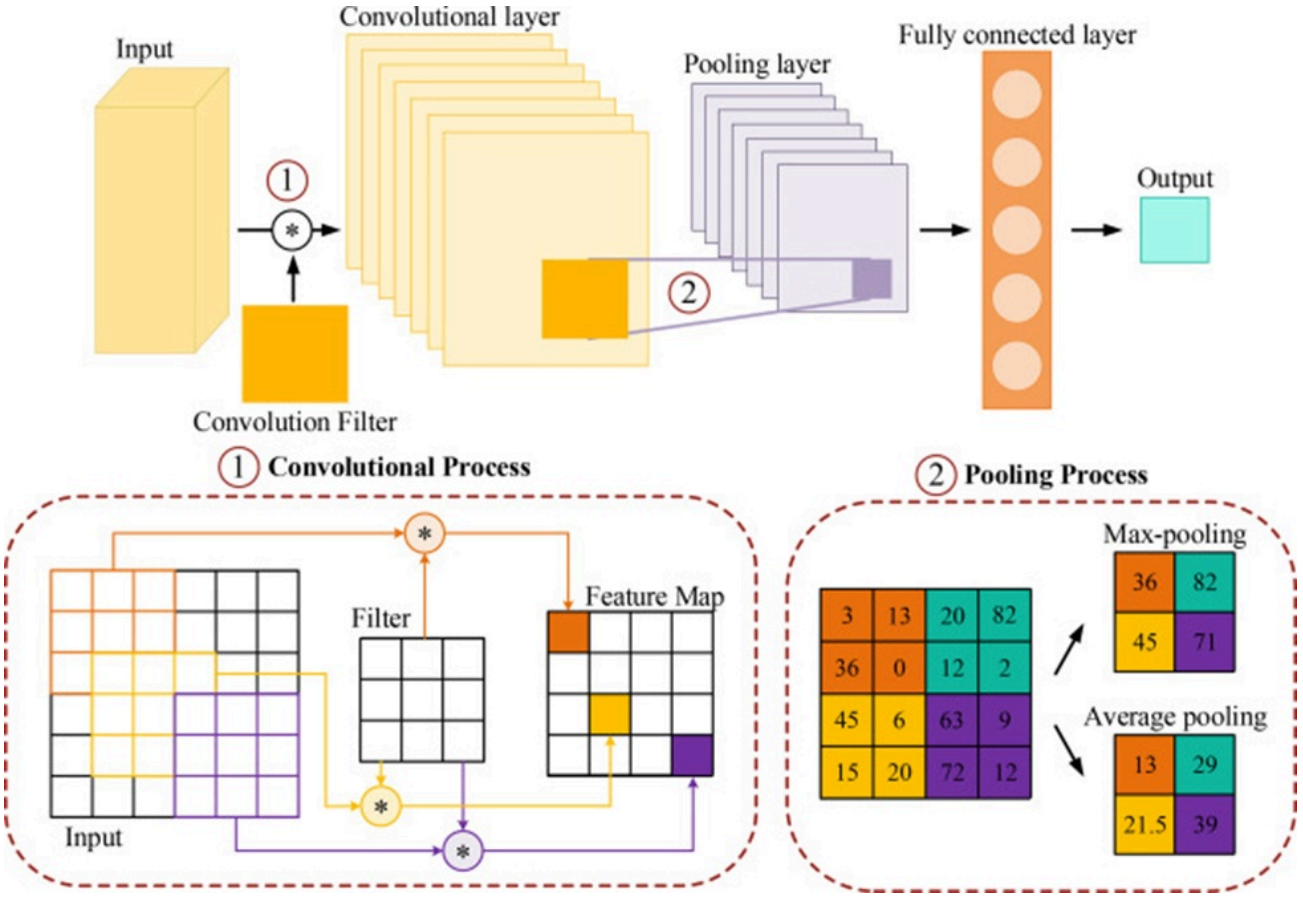


Figure 5. Concept of convolutional neural network.

Fully connected layer

After spatial reduction, dense layers process the data by executing matrix multiplications against a weight array and introducing a global bias. This fully connected architecture also flattens and resizes the internal vectors to align precisely with the target dimensional space of the model's outputs.

Output layer

The terminating layer delivers the final prognostication and initiates backward error propagation across the preceding network to adjust the weights and biases, aimed at minimizing the regression loss. A halved mean squared error serves as the core training criterion, formalized in **Eq. 7.** To sweep for the most effective hyperparameters across all segments, a systematic grid search is applied, targeting the maximization of validation R^2 within the operational bounds outlined in **Table 2.**

$$HMSE = \frac{1}{2N}\sum_{t=1}^{N}\left(y_{actual,t} - y_{predict,t}\right)^2 \quad (7)$$

Table 2. Searching domains for hyperparameter tuning.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Layer category	Hyperparameter	Search range/Value
Convolutional layer	Filter size	[5, 20]
	Number of filters	[10, 60]
Pooling layer	Pooling window size	[1, 6]
	Pooling method	[Average, Max]
Fully connected layer	Number of neurons	[1, 60]
Training settings	L2 regularization coefficient	[0, 1]
	Learning rate	[0.25, 1]
	Learning rate decay period	[20, 100]
	Learning rate decay factor	[0.5, 1]
	Maximum number of epochs	[100, 500]
	Validation frequency	10
	Mini-batch size	120
	Weight update optimization algorithm	[Adam, SGD]

Transfer learning procedure

Knowledge transfer functions as a powerful instrument to circumvent initial weight assignment issues while enhancing the stability and variance-resistance of architectures handling high-dimensional arrays [37]. This study adopts a parameter-driven transfer learning paradigm to eliminate the necessity for massive secondary data gathering and drastically curtail computational training durations [38].

Our empirical analysis benchmarks eight individual computational paradigms for tracking energy metrics: a regression tree (RT), a feed-forward neural network (FFNN), a recurrent neural network (RNN), an LSTM network, a vanilla single-channel CNN, a dual-channel CNN (2C-CNN), a triple-channel CNN lacking transfer capabilities (3C-CNN), and the newly introduced TLCNN.

Synthetic data imputation via linear interpolation and zero-order hold (ZOH) configurations are likewise evaluated as baseline control setups. These two paradigms represent the most elementary data restoration mechanisms available. The ZOH method bridges missing records by sustaining the last valid data state until a new reading is registered. Conversely, the interpolation method presumes a straight-line trend between successive known data boundaries to approximate intervening values. Nonetheless, these imputation routines inject significant computational overhead into the pipeline, rendering them incompatible with real-time operational deployment where future boundary coordinates remain unknown.

The 2C-CNN configuration deploys a dual-stream architecture, whereas the 3C-CNN framework expands this to three parallel pathways. For both models, the baseline channel mirrors the structural design of the basic CNN. In contrast, the secondary processing stream of the 2C-CNN and the tertiary pathway of the 3C-CNN rely on distinct hyperparameter matrices derived from an expanded grid search routine.

In contrast, the proposed TLCNN paradigm initializes its first two structural pathways using the pre-conditioned weight distributions harvested from the 2C-CNN framework, dedicating a separate training phase exclusively to optimizing its third pathway. This parallelized system secures heightened optimization stability by partitioning the feature-extraction

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

burden across separate convolutional streams, enabling the third pipeline to map unique signal dynamics from the target dataset.

Selective layer immobilization is applied across the primary channels to compel the newly added third stream to assimilate novel trends contained within the incoming data. By locking the internal tracking weights of the pre-trained modules, the model anchors the feature knowledge acquired from the initial data rate, while simultaneously adapting the newly appended convolutional stream to accommodate modified sampling frequencies. This architectural assembly and the associated transfer learning workflow are illustrated in **Figure 6.**

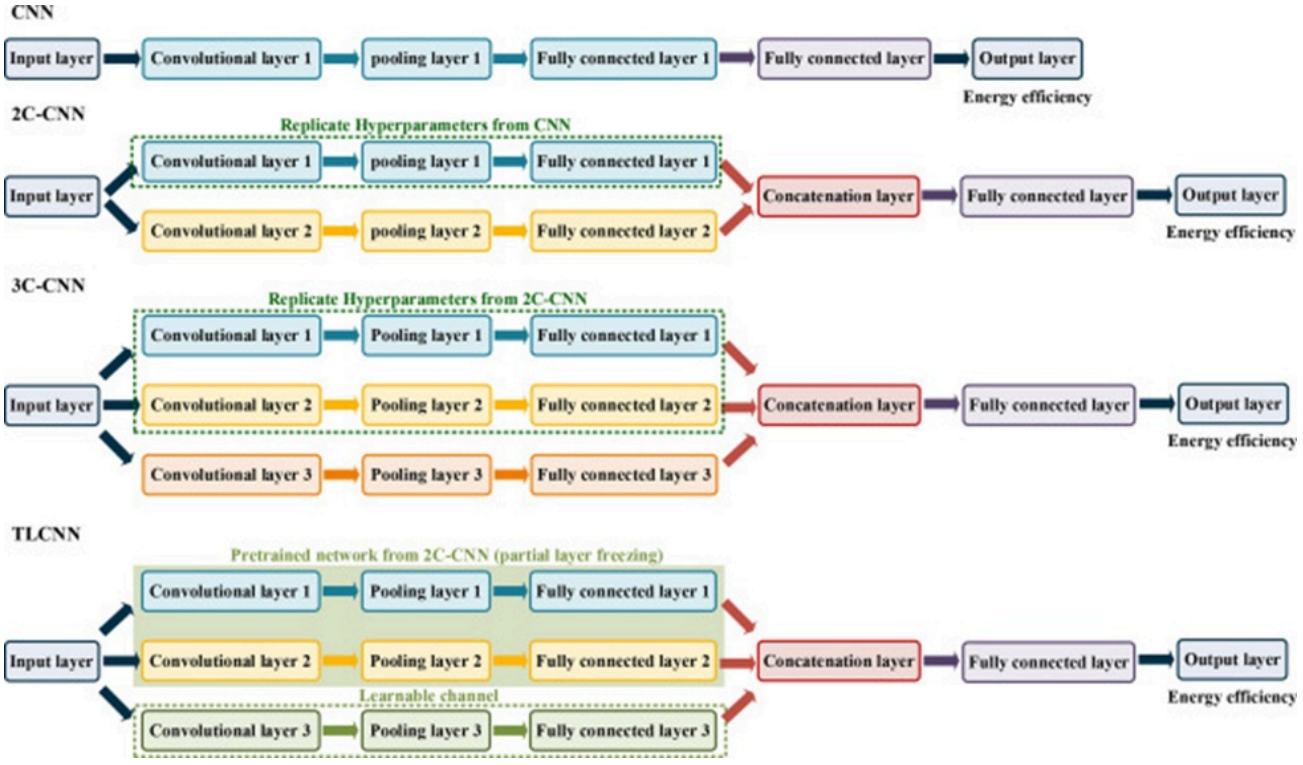


Figure 6. CNN formulation and transfer learning procedure.

Model testing

Ultimately, all models undergo evaluation through software-in-the-loop testing for real-time applications, a reproducibility assessment to verify model consistency and reliability ratios, and a sliding test to observe time-scale variations post-training. These are conducted using R-squared (R2), root mean squared error (RMSE), and mean absolute percentage error (MAPE) as performance metrics, presented in **Eqs. 8–10**. Detailed information regarding the testing protocol will be provided in the results and discussion chapter.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (8)$$

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2} \quad (9)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}}{y_i}\right| \times 100\% \quad (10)$$

Where yi denotes the actual output for the i th sample, $\hat{y}$ represents the predicted output, and $\bar{y}$ signifies the average of the original output. The R2 metric serves as a measure of model fitness, indicating the fraction of variance in energy efficiency that can be attributed to input variables through the model. Furthermore, RMSE and MAPE quantify the discrepancy between predicted and actual outcomes. While RMSE accounts for the absolute magnitude of errors, offering a more exhaustive perspective on model efficacy than R2, MAPE is expressed as a percentage, which makes the model's error less susceptible to extreme values than RMSE.

Vinyl Chloride Monomer Process

The VCM production process comprises five distinct stages: direct chlorination, oxychlorination, ethylene dichloride (EDC) purification, EDC pyrolysis, and VCM purification. **Figure 7** illustrates the general layout of the VCM process and its constituent sections. Operations commence by introducing ethylene and chlorine gas into the direct chlorination unit, where an addition reaction yields EDC. During the oxychlorination phase, ethylene undergoes a gas-phase reaction with oxygen and hydrogen chloride to produce EDC, which is subsequently neutralized using a sodium hydroxide solution within a caustic scrubber. The resulting output is directed to a distillation column to isolate EDC from water. This purified EDC is then consolidated from three streams: EDC originating from direct chlorination, EDC from oxychlorination, and unreacted EDC recovered from the VCM purification stage. Within the EDC purification section, the combined feed is refined to eliminate trace impurities, including water, ethyl chloride, and trichloroethane (TCE). The primary column extracts light fractions, low-boiling-point contaminants, and water, whereas the secondary column removes heavier components such as TCE. Purified EDC is then fed into a thermal cracker equipped with extensive tubular coils to synthesize VCM and associated by-products, operating at temperatures between 480 and 530 °C and pressures spanning 607.95 to 3546.37 kPa. These hot gases are subsequently cooled via a series of condensers and routed to a hydrogen chloride (HCl) column for separation, where unreacted HCl is recovered and returned to the oxychlorination stage as feedstock for EDC synthesis. A secondary column partitions the VCM product from unreacted EDC, which is recycled to the EDC purification unit for further cracking.

In summary, VCM production involves several conversion phases, including chlorination, cracking, and various purification procedures. These processes require high temperatures and pressures, which demand substantial energy inputs to sustain productivity and meet utility requirements. Consequently, the VCM manufacturing sector is characterized by high energy consumption.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

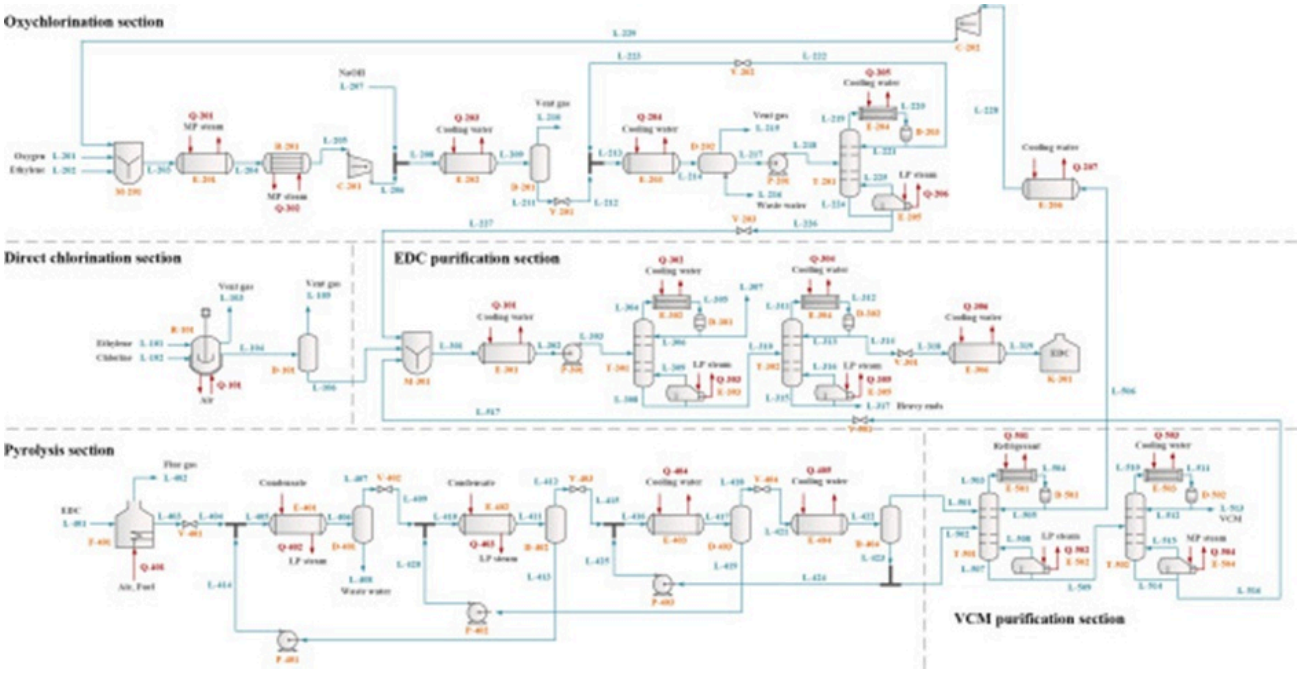


Figure 7. Overall process of VCM production.

Multirate information

The data used in this research were obtained via co-simulation between MATLAB and the UniSim Design Suite, using simulation parameters derived from the literature [39]. In this investigation, ethylene input fluctuates between 235,000 and 410,000 kg/day (comprising a minimum capacity of 160,000 kg/day from direct chlorination and 75,000 kg/day from the oxychlorination unit). At the same time, fresh EDC feeds to the pyrolysis unit vary from 750,000 to 1,800,000 kg/day, with a random distribution within a ±10% range. The dataset consists of 2000 samples generated at 1-minute intervals within the UniSim environment. Finally, data are aggregated across multiple time scales using MATLAB, in accordance with **Table 3,** at distinct rates of 1, 5, and 10 min, assuming constant sampling frequencies during the training phase. The VCM process involves 97 measurable variables; an illustration of multirate information is provided in **Figure 8a**. Because transitions between consecutive records of low-sampling-rate variables are obscured, the resulting data lack high-frequency dynamics, complicating the identification of underlying patterns. Data alignment yields three distinct configurations based on the available information: 1 min, 1–5 min, and 1–5–10 min, as depicted in **Figure 8b.**

Table 3. Description of process variables in the VCM process and sampling intervals.

No	Process variable	Interval	No	Process variable	Interval
1	Temperature of EDC at bottom D-101	1 min	50	CW feed rate of E-304	1 min
2	Mass flow rate at bottom R-101	1 min	51	CW feed rate of E-301	5 min
3	Mass flow rate at top D-101	5 min	52	CW feed rate of E-306	10 min
4	Mass flow rate of EDC at bottom D-101	1 min	53	LP steam feed rate of E-303	1 min
5	Mass flow rate of chlorine feed to R-101	1 min	54	LP steam feed rate of E-305	1 min
6	Mass flow rate of ethylene feed to R-101	1 min	55	Temperature of EDC feed to F-401 (1)	1 min

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

7	Air feed rate of R-101	5 min	56	Temperature of EDC feed to F-401 (2)	1 min
8	LP steam feed rate of R-101	1 min	57	Temperature at L-403	1 min
9	Temperature of recycled HCl at L-228	5 min	58	Temperature of recycled EDC at D-402	1 min
10	Temperature of recycled HCl at L-229	1 min	59	Temperature at D-401 inlet	1 min
11	Temperature at C-201 outlet	10 min	60	Temperature at top D-401	1 min
12	Temperature at P-201 outlet	5 min	61	Temperature of recycled EDC D-403	1 min
13	Temperature at bottom D-201	10 min	62	Temperature at D-402 inlet	1 min
14	Temperature of recycled water at L-223	10 min	63	Temperature of recycled water from D-404	1 min
15	Temperature of decanter D-202	1 min	64	Temperature at D-403 inlet	1 min
16	Temperature at top T-201	1 min	65	Temperature at D-404 inlet	1 min
17	Temperature at bottom T-201	1 min	66	Mass flow rate of EDC feed to F-401	1 min
18	Mass flow rate of ethylene feed to M-201	1 min	67	Mass flow rate at top D-401	1 min
19	Mass flow rate of oxygen feed to M-201	1 min	68	Mass rate of drainage at bottom D-401	10 min
20	Mass flow rate of EDC at bottom T-201	1 min	69	Mass flow rate at top D-402	5 min
21	Mass flow rate at L-208	10 min	70	Mass rate of recycled at bottom D-402	5 min
22	Mass flow rate of recycled HCl at L-229	1 min	71	Mass flow rate at top D-403	1 min
23	Mass flow rate at D-202 inlet	5 min	72	Mass flow rate at top D-404	5 min
24	Mass flow rate of NaOH at L-207	1 min	73	Mass rate of recycled at bottom D-403	5 min
25	Mass flow rate at top D-201	5 min	74	Mass rate of recycled at bottom D-404	1 min
26	Mass flow rate at bottom D-202	5 min	75	Mass flow rate at P-403 outlet	1 min
27	Mass flow rate at P-201 inlet	1 min	76	Mass flow rate at D-402 inlet	5 min
28	Mass flow rate at top T-201	1 min	77	Mass flow rate at D-401 inlet	1 min
29	CW feed rate of E-202	1 min	78	CW feed rate of E-403	10 min
30	CW feed rate of E-203	5 min	79	CW feed rate of E-404	1 min
31	CW feed rate of E-204	1 min	80	Fuel feed rate of F-401	1 min
32	LP steam feed rate of E-205	1 min	81	LP steam feed rate of E-401	1 min
33	MP steam feed rate of E-201	1 min	82	LP steam feed rate of E-402	5 min
34	MP steam feed rate of R-201	5 min	83	Temperature at top T-501	1 min
35	CW feed rate of E-206	1 min	84	Temperature at bottom T-501	1 min
36	Power input of C-202	5 min	85	Temperature at top T-502	1 min
37	Temperature of EDC feed to E-301	10 min	86	Temperature at bottom T-502	1 min
38	Temperature of EDC feed to T-301	5 min	87	Temperature of recycled EDC at L-517	1 min
39	Temperature at top T-301	1 min	88	Mass flow rate at L-501	1 min

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

40	Temperature at bottom T-301	1 min	89	Mass flow rate at L-502	1 min
41	Temperature at top T-302	1 min	90	Mass flow rate at T-502 inlet	1 min
42	Temperature at bottom T-302	1 min	91	VCM mass flow rate	1 min
43	Temperature of purified EDC at K-301	10 min	92	Mass rate of recycled at bottom T-502	1 min
44	Mass flow rate of EDC feed to E-301	1 min	93	Mass rate of recycled at top T-501	1 min
45	Mass flow rate at top T-301	1 min	94	CW feed rate of E-503	1 min
46	Mass flow rate at bottom T-301	1 min	95	LP steam feed rate of E-502	1 min
47	Mass flow rate at top T-302	1 min	96	MP steam feed rate of E-504	1 min
48	Mass flow rate at bottom T-302	1 min	97	Refrigerant feed rate of E-501	1 min
49	CW feed rate of E-302	1 min	–	–	–

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

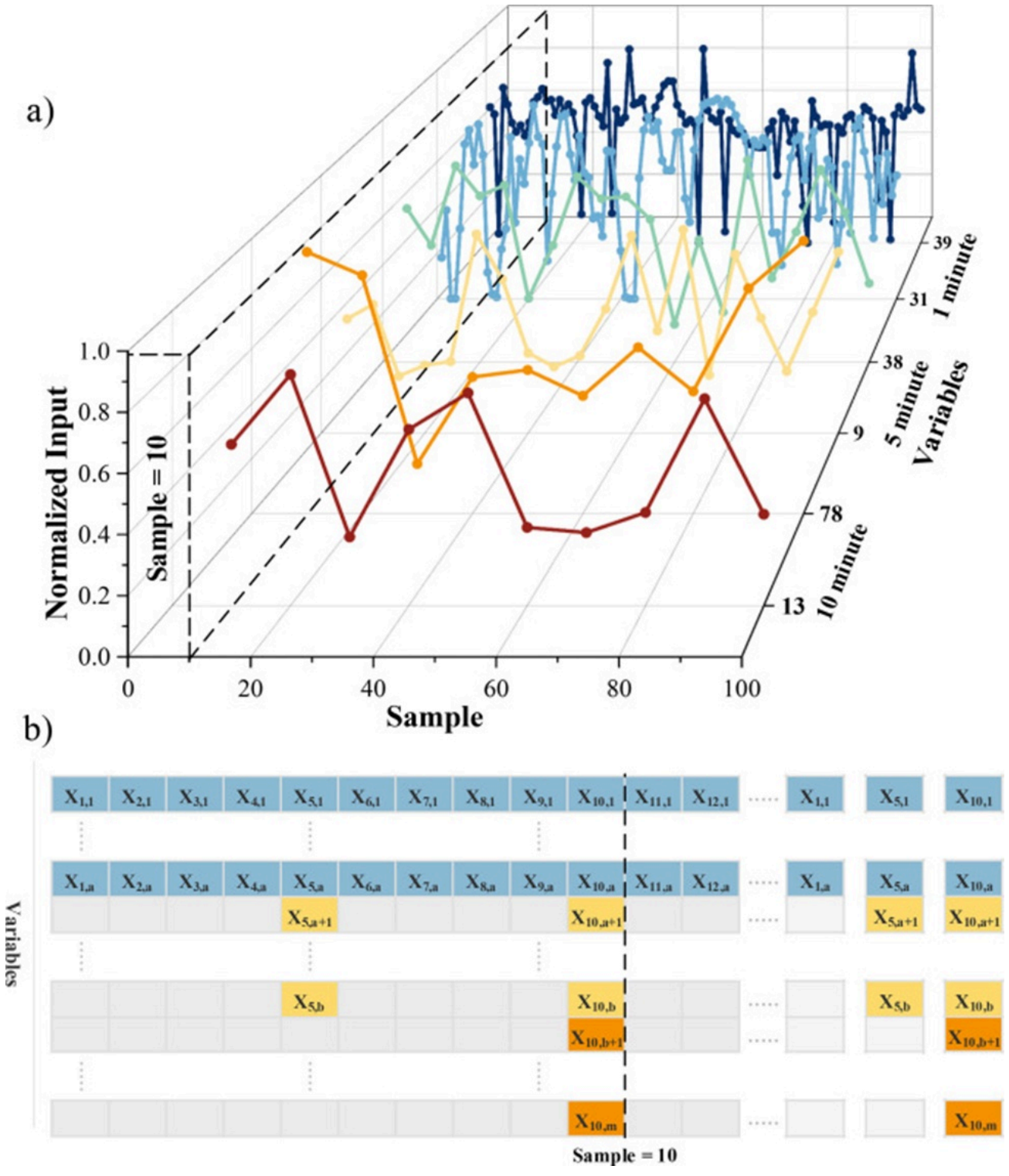


Figure 8. Multirate sampling interval including a) example of data and b) combination of input.

To guarantee the precision and generalizability of the deep learning framework, the dataset is partitioned into training and validation subsets. The training set comprises 80% of the total data, while the remaining 20% is reserved for validation. Throughout the training procedure, holdout validation is implemented with a validation frequency of 2 (training iterations per validation iteration); this assesses model accuracy using unseen data throughout the training cycle, thereby mitigating the risk of overfitting.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Specific energy consumption

Metrics such as energy demand, energy intensity, and thermal efficiency are frequently utilized to characterize energy efficiency [34]. However, evaluating energy efficiency performance across varying plant loads proves difficult since these traditional parameters account solely for energy consumption and conversion efficiency [40]. An alternative and valuable metric for evaluating energy efficiency is specific energy consumption (SEC) [41]. SEC represents the amount of energy needed to manufacture a specific product relative to the total production capacity, and it can be computed via **Eq. 11.**

$$SEC = \frac{\text{EnergyUsed}}{\text{AmountofProductProduced}} \quad (11)$$

Based on its definition, SEC standardizes energy demand per functional unit of a product, enabling straightforward comparisons across different production capacities. Furthermore, SEC serves as an effective benchmarking tool across various scales—spanning process, regional, and global levels—and facilitates comparisons among different industries manufacturing comparable goods.

Results and Discussion

Feature importance score

To identify the most critical process variables, the entire set of 97 variables was ordered from highest to lowest significance based on their FI score. During this initial phase, any variable presenting an FI score of zero was removed. Consequently, the 40 variables that remained were subsequently applied in the forward stepwise feature selection process. **Figure 9** illustrates the outcomes of this forward stepwise selection utilizing the FI rankings during training. Based on the documented error and accuracy metrics, it is evident that the highest training R2 and the lowest RMSE are achieved when exactly 14 predictor variables are used. **Table 4** details these essential variables alongside their corresponding FI scores. Ultimately, the forward stepwise feature selection validates that the final set consists of the 14 variables possessing the highest FI scores, perfectly aligning with their initial feature importance rankings.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

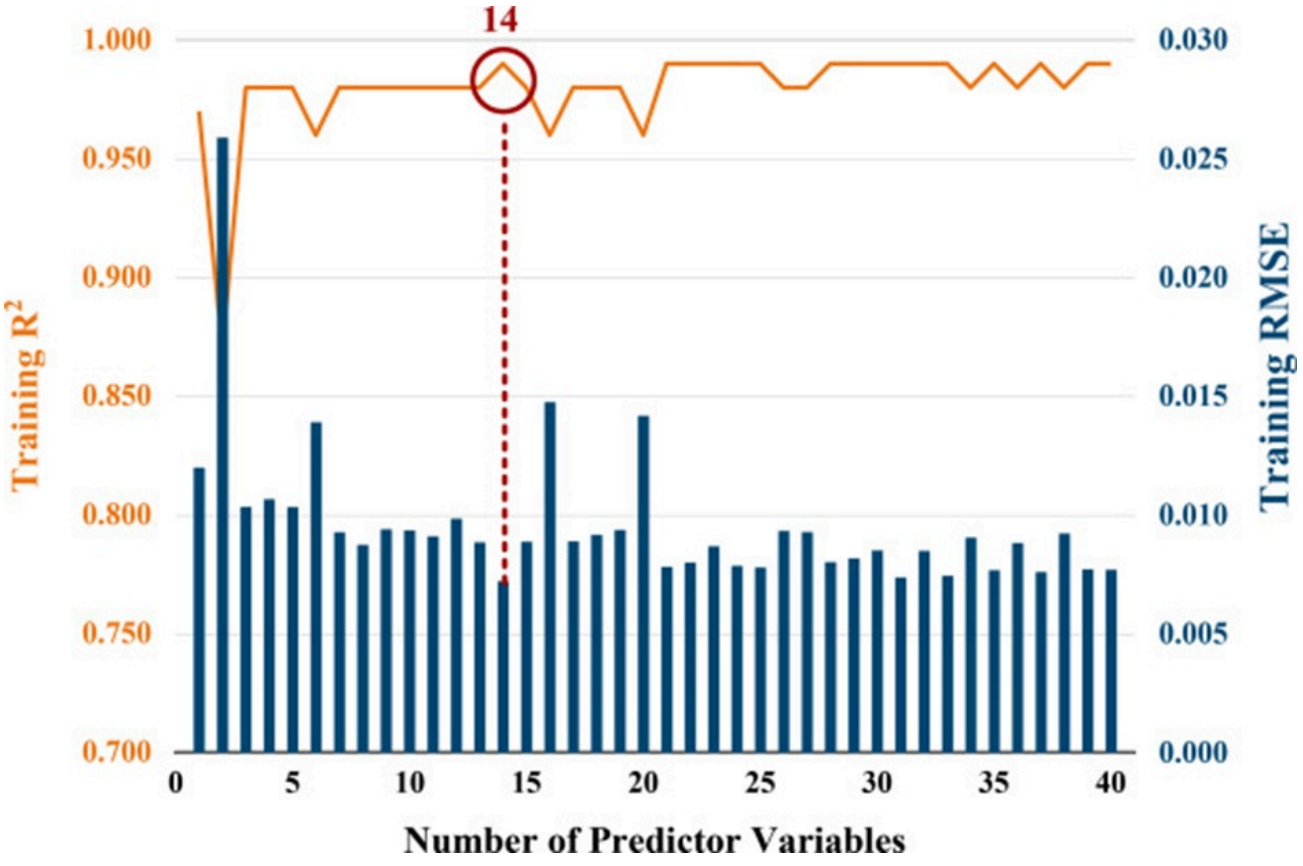


Figure 9. R2 and RMSE from model training of forward stepwise feature selection.

Table 4. Significant input variables for model development.

No.	Process variable	Sampling interval	FI score
63	Temperature of recycled EDC from D-404	1 min	15.15%
1	Temperature measured at the bottom of D-101 for EDC	1 min	14.61%
84	Temperature at the bottom section of T-501	1 min	10.15%
23	Inlet mass flow rate to D-202	5 min	9.51%
9	Temperature of recycled HCl at L-228	5 min	8.31%
47	Mass flow rate at the top of T-302	1 min	8.29%
46	Mass flow rate at the bottom of T-301	1 min	7.17%
78	Cooling-water (CW) feed rate to E-403	10 min	7.03%
31	Cooling-water (CW) feed rate to E-204	1 min	6.39%
85	Temperature at the top of T-502	1 min	3.05%
10	Temperature of recycled HCl at L-229	1 min	2.53%
13	Temperature at the bottom of D-201	10 min	1.94%
39	Temperature at the top of T-301	1 min	1.78%

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

38	Temperature of EDC entering T-301	5 min	1.04%
–	Remaining variables (aggregate of 83 variables)	–	3.05%

Hyperparameters and speed indicator

Table 5 presents the optimal hyperparameters derived from grid-search optimization that yield the highest training R2 for every CNN model utilized in this research. The findings indicate that both the TLCNN and 3C CNN frameworks required 14 hyperparameters. The primary distinction between the TLCNN and 3C CNN is that the TLCNN imports both hyperparameters and learnable parameters. In contrast, the 3C CNN merely copies hyperparameters from an earlier iteration of the multi-channel models. **Figure 10** illustrates the balance between accuracy and training duration across all assessed CNN models.

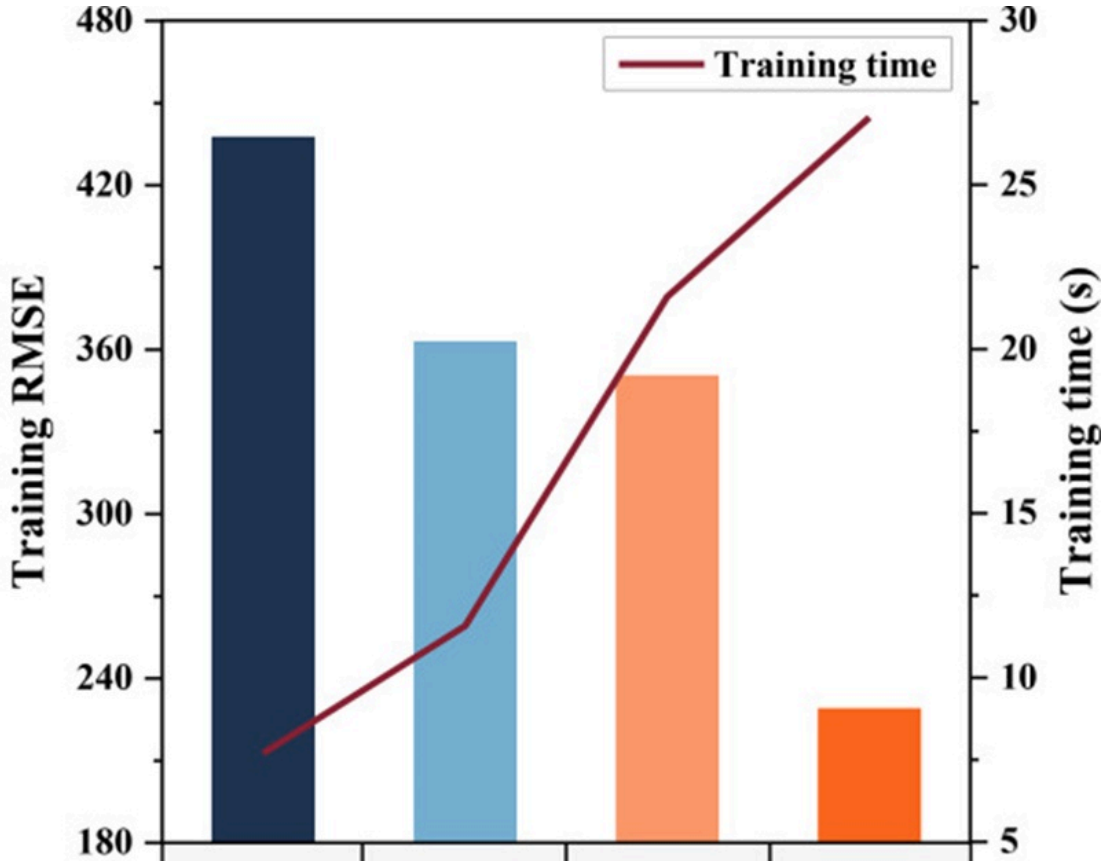
Given that the prediction speeds for all models fall under one second (which represents the most common sampling interval in industrial measurements), every model is suitable for real-time application. The TLCNN requires a training duration of 28 s. Because the actual sampling rate occurs on a minute-by-minute basis, these results demonstrate that the models can complete their retraining cycles before the next sample is acquired. Consequently, the 2C CNN and 3C CNN architectures require more computational time than a standard CNN model. While the TLCNN demands the longest training period, it simultaneously achieves the lowest training RMSE relative to the other learning algorithms. Furthermore, the TLCNN forecasts future samples more rapidly than the 3C CNN, despite both possessing an identical number of channels.

Table 5. Optimal hyperparameters for CNN models.

Tuning	Hyperparameters	CNN	2CCNN	3CCNN	TLCNN
1	Size of filters_1	10	10	10	10
	Number of filters_1	50	50	50	50
	Size of PR_1	3	3	3	3
	Number of FC nodes_1	50	50	50	50
2	Size of filters_2	–	10	10	10
	Number of filters_2	–	50	50	50
	Size of PR_2	–	3	3	3
	Number of FC nodes_2	–	1	1	1
3	Size of filters_3	–	–	15	15
	Number of filters_3	–	–	10	50
	Size of PR_3	–	–	5	5
	Number of FC nodes_3	–	–	1	40
Global	Learning rate	0.75	0.75	0.75	0.75

Note: PR = pooling regions, FC = fully connected.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press



	CNN	2C-CNN	3C-CNN	TLCNN
Training RMSE	437.334	362.702	350.351	228.738
Training Time (s)	7.697	11.577	21.588	27.049
Prediction Speed (obs/s)	6,413	6,105	4,380	5,062

Figure 10. Comparison of training RMSE and training time using different CNN models.

Energy efficiency prediction testing performance

This segment presents the online testing outcomes regarding the energy efficiency predictions of each model. To evaluate real-time operational performance, the model was integrated into the UniSim Design platform. Average performance metrics were established by running the models through 20 iterations, utilizing 400 samples per iteration, with the final findings detailed in **Table 6.**

An evaluation of the FFNN, RNN, LSTM, and CNN models reveals that the CNN generally demonstrates superior accuracy when processing predictors across multiple time scales. As a result, the CNN was chosen as the forecasting model for data reconstructed via interpolation and zero-order hold methods. A basic CNN achieves an accuracy of 0.929;

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

however, when integrated with zero-order hold and interpolation, the accuracy shifts to 0.907 and 0.915, respectively. Demonstrating superior predictive capability, the TLCNN surpasses all other evaluated regression models, achieving an R2 of 0.970, an RMSE of 232.862, and a MAPE of 0.784. This superiority is largely because reconstruction methods like interpolation and zero-order hold rely on inherent assumptions regarding the baseline dataset—specifically, a zero-order hold estimates missing samples by presuming that the information remains unchanged. In contrast, interpolation postulates that unrecorded values situated between two known points follow a single uniform relationship. Consequently, interpolation inaccuracies and the loss of data from shifting patterns can degrade the predictive accuracy generated by neural networks.

Compared to a standard CNN, the introduced model improves SEC tracking capability by 43.13%, surpassing the gains observed in the 2D-CNN (17.06%) and 3D-CNN (19.88%). These findings indicate that merely adding channels to a model can lead to diminishing performance gains and increase the likelihood of overfitting. By employing transfer learning, the bias and weight factors of the pre-trained channel are locked in, compelling the architecture to identify residual data patterns through the untrained channel. This approach effectively enhances both the overall performance and the generalization capability of the model.

Table 6. Software-in-loop testing results.

Dataset	Model	R^2	RMSE	MAPE
Training	FFNN	0.741	899.359	2.546
	RNN	0.891	433.708	1.392
	LSTM	0.904	498.773	1.511
	CNN	0.929	437.334	1.482
	CNN (Zero-orderhold)	0.851	483.642	1.667
	CNN (Interpolation)	0.871	452.174	1.551
	2CCNN	0.951	362.702	1.227
	3CCNN	0.956	350.351	1.183
	TLCNN	0.971	228.738	0.760
Validation	FFNN	0.832	848.018	1.846
	RNN	0.813	828.486	1.992
	LSTM	0.921	457.401	1.538
	CNN	0.927	436.898	1.474
	CNN (Zero-order hold)	0.868	487.621	1.678
	CNN (Interpolation)	0.872	475.678	1.631
	2CCNN	0.942	388.743	1.288
	3CCNN	0.954	355.058	1.185
	TLCNN	0.962	262.518	0.803
Testing	FFNN	0.771	869.828	2.315
	RNN	0.617	1205.701	3.501

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

LSTM	0.919	462.329	1.650
CNN	0.929	392.484	1.262
CNN (Zero-order hold)	0.907	471.861	1.686
CNN (Interpolation)	0.915	443.578	1.566
2CCNN	0.945	319.277	1.094
3CCNN	0.948	311.666	1.039
TLCNN	0.970	232.862	0.784

Figure 11 provides a comparison between the measured SEC and the SEC values forecasted by the various CNN-based models applied to the VCM process. The findings indicate that the TLCNN can accommodate a broad spectrum of SEC values, whereas the standard CNN, 2C-CNN, and 3C-CNN models are effective only when the true SEC values range between 22,000 and 24,000 kJ/kg. Incorporating transfer learning successfully establishes a stable equilibrium between variance and bias, a balance that competing prediction models frequently struggle to maintain in real-time implementations. Within this architecture, the pre-trained channel functions primarily to extract features. In contrast, the final channel acts to fine-tune the network for the specific task, thereby expanding the operational domain of the model.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

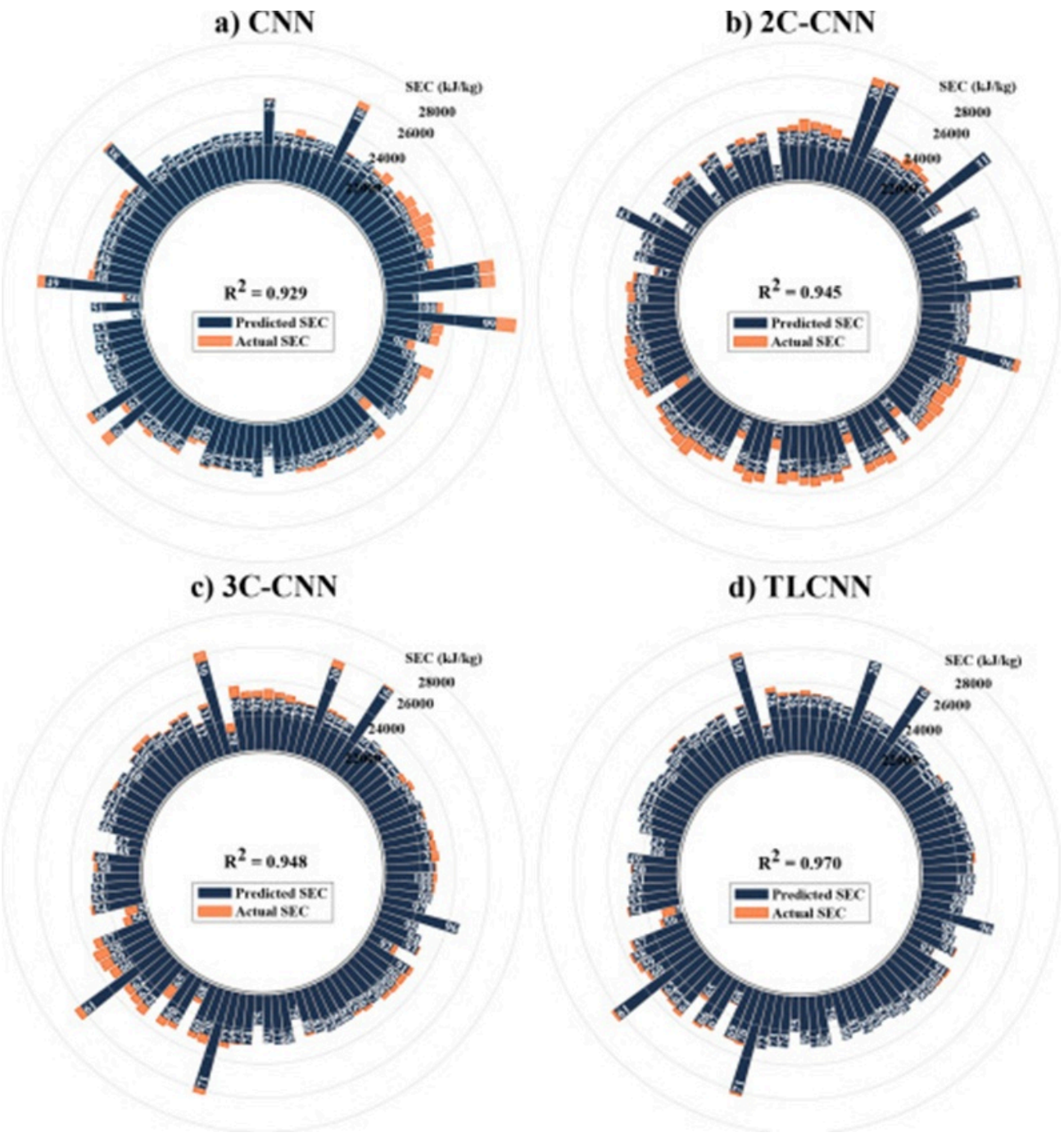


Figure 11. Example of comparison between actual and predicted SEC of all CNN-based models.

Reproducibility analysis

When deploying artificial intelligence for advanced predictive analytics, prioritizing accuracy alone is insufficient; ensuring model reproducibility is equally vital. This section details the reproducibility assessment to confirm that the generated outcomes can be consistently replicated. To evaluate potential inconsistencies, every model underwent 200 separate retraining iterations. During these trials, the optimal hyperparameters were kept constant, but the models were fed with freshly randomized datasets presented in varying sequences. **Figure 12** visualizes the outcomes of this rigorous testing

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

protocol. Adaptive architectures, particularly the TLCNN, demonstrate markedly superior replicability when juxtaposed against standard, single-training setups. Specifically, excluding anomalous data points, the testing R2 for the TLCNN fluctuates strictly between 0.929 and 0.992.

In contrast, the standard CNN's performance is confined to a lower bracket of 0.874 to 0.933. Crucially, while the inclusion of multiple channels in the 2C-CNN and 3C-CNN configurations successfully elevates the peak testing R2, it simultaneously causes a severe deterioration in their lowest prediction thresholds—plummeting to 0.854 and 0.869, respectively. This degradation substantially heightens the risk of failure during the training phase of model creation. Integrating additional channels complicates the optimization pathway, making convergence toward a viable solution much harder. To circumvent these weight initialization bottlenecks, model parameter-based transfer learning becomes an essential intervention. By acting as a predetermined anchor, the pre-trained network effectively narrows down the optimal zones for weights and biases, while simultaneously cutting down the volume of parameters that must be learned from scratch.

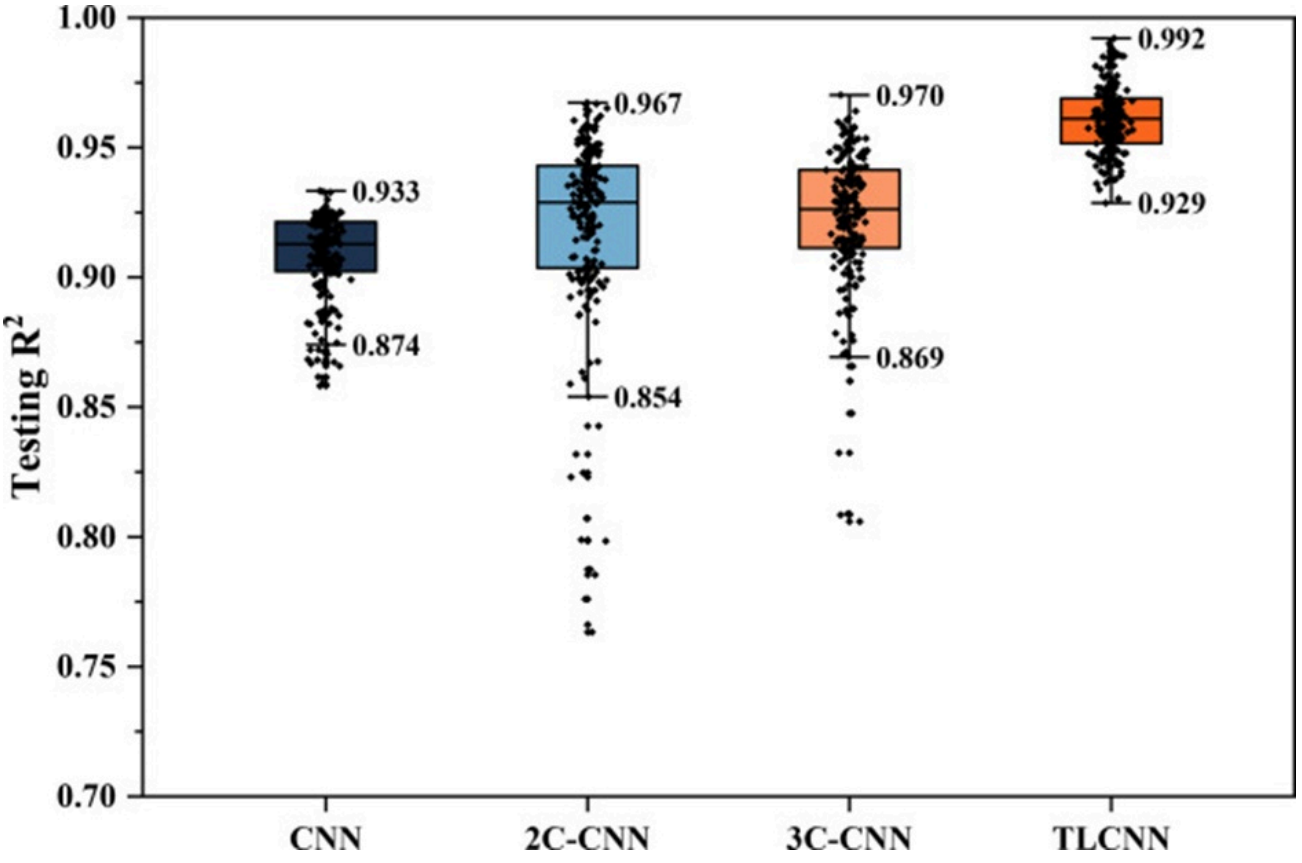


Figure 12. Boxplot of testing R2 values of CNN models.

Sliding test

In real-world applications, a fully trained model is transitioned into a deployment phase to track operational metrics using novel incoming datasets. Conventionally, the testing sets analyzed in literature mirror the structural cadence of the training sets (1–5–10 min). However, when tackling multirate prediction challenges, it becomes critical to scrutinize the system's robustness against temporal shifts, specifically when the incoming data rhythm deviates from what the model originally learned. **Figure 13** highlights multiple potential scenarios where sampling timeframes shift within industrial settings. The scenarios designated as Testing 1 and 2 highlight the impact of system latency, illustrating sampling frames that are either synchronized with the least common multiplier (1–6–12 min) or disjointed from it (1–6–11 min). Conversely,

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Testing 3 and 4 portray situations where specific variables alter their transmission rates—either accelerating in frequency or vanishing momentarily amidst maintenance protocols.

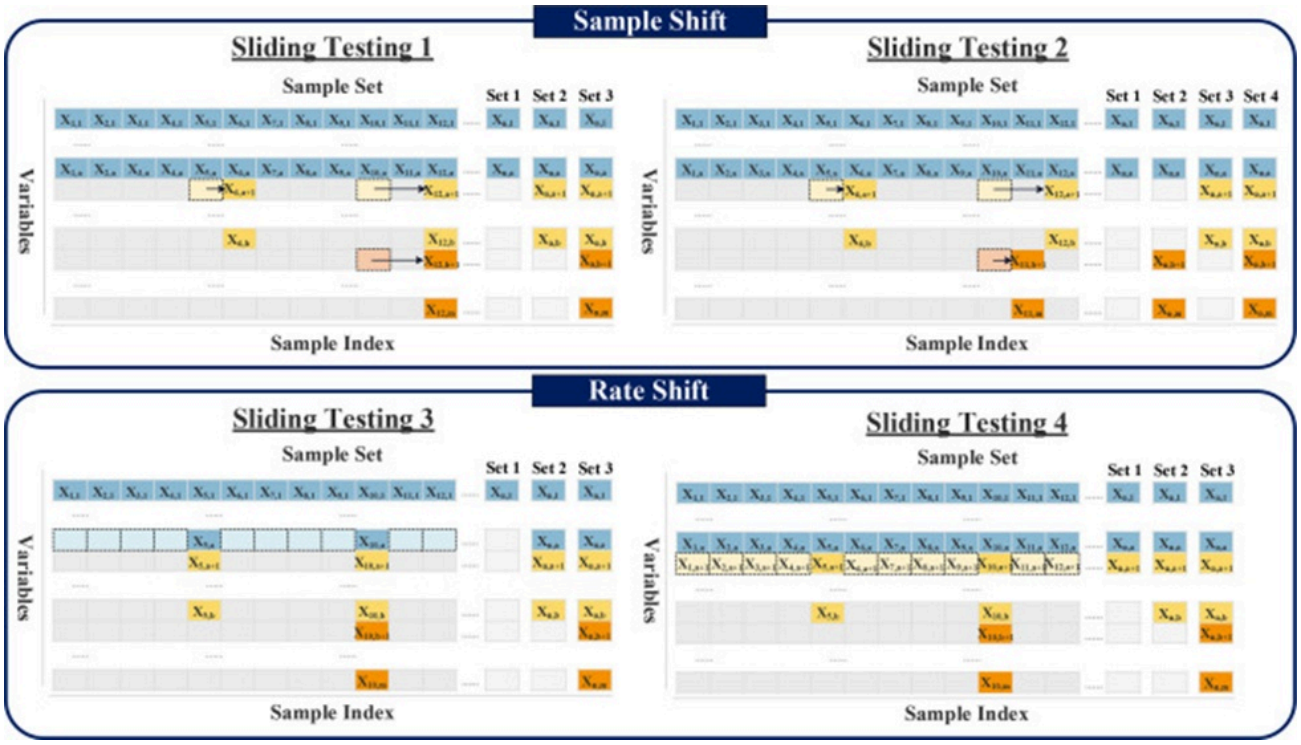


Figure 13. Sliding testing scenario.

Table 7 outlines how different models react across a 400-sample test window subjected to these altered time scales. The empirical evidence underscores that exceptional performance is only sustained universally during Testing 1. In this isolated case, the three potential subgroupings of the testing input (1, 1–6, 1–6–12 min) perfectly match the combinations seen during training (1, 1–5, 1–5–10 min). Yet, the moment a novel structural pattern—one entirely unseen during the training phase—manifests in the testing environment, every model suffers a pronounced drop in efficacy. This is where transfer learning demonstrates profound utility; the fine-tuning mechanism embedded within the final channel grants the architecture the flexibility to decipher nuances in the unfamiliar input. By locking the primary learnable parameters while allowing the final channel to adjust iteratively, the network's overall adaptability surges. Consequently, the pre-trained segments recycle wisdom gleaned from historical input patterns, leaving the terminal channel free to calibrate itself against the idiosyncrasies of the newly introduced sampling cadence.

Table 7. Result of CNN-based model on a sliding test.

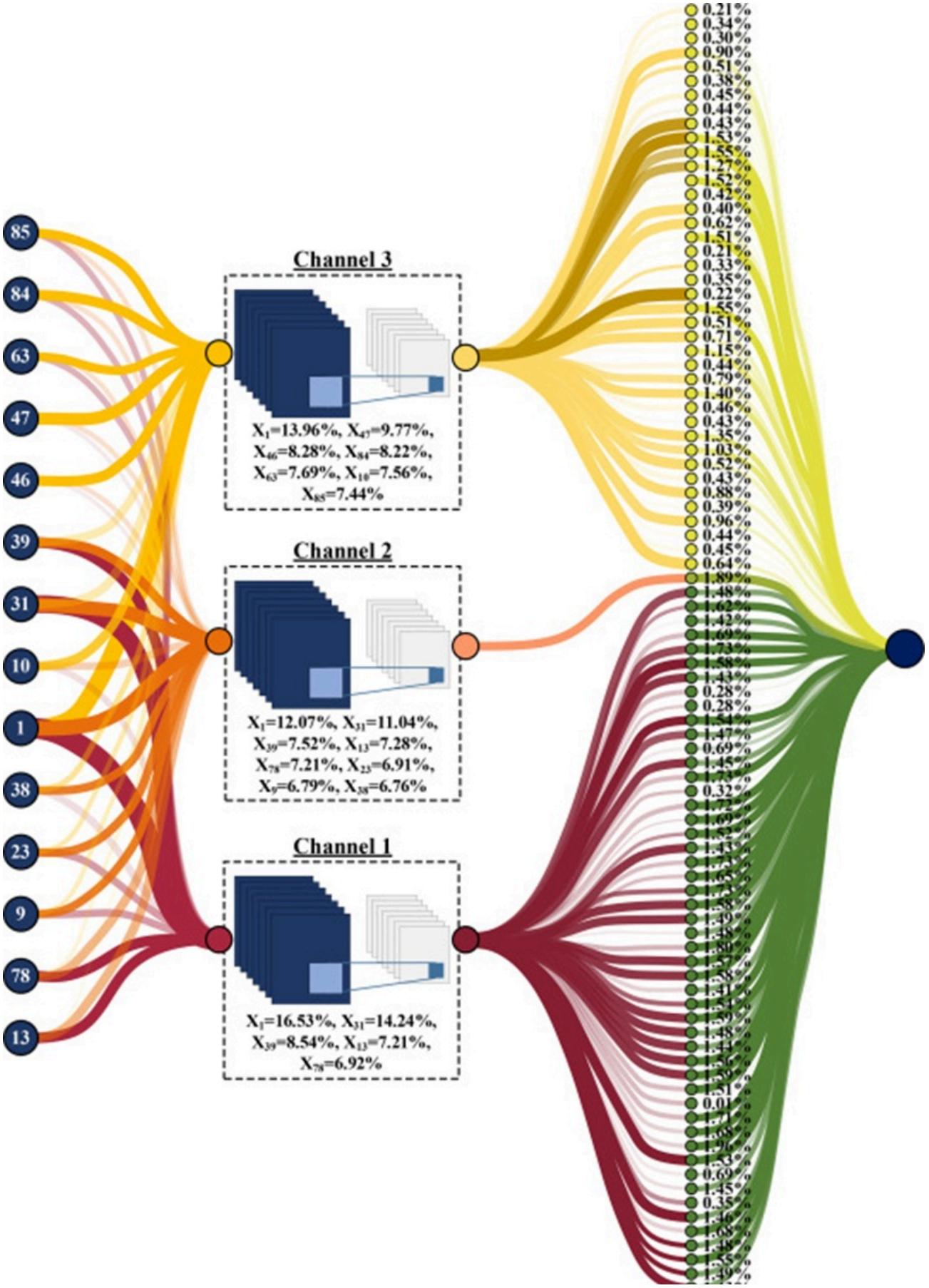
Scenario	Dataset	Model	Testing R^2	RMSE	MAPE
Sample shift	Test 1	CNN	0.935	334.688	1.152
		2C-CNN	0.931	344.871	1.187
		3C-CNN	0.933	340.815	1.181
		TLCNN	0.954	271.384	0.919
	Test 2	CNN	0.752	820.529	1.652
		2C-CNN	0.662	918.777	2.091

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

		3C-CNN	0.779	653.652	1.779
		TLCNN	0.938	337.797	1.071
Variable shift	Test 3	CNN	0.392	2137.721	4.163
		2C-CNN	0.398	2077.245	4.048
		3C-CNN	0.464	1767.844	3.755
		TLCNN	0.678	985.981	2.671
	Test 4	CNN	0.471	2678.653	9.565
		2C-CNN	0.462	2598.033	9.157
		3C-CNN	0.505	2257.201	8.161
		TLCNN	0.815	595.259	2.024

Applying the principle of feature attribution derived from weights [42], **Figure 14** depicts both the weight distribution allocated by individual TLCNN filters and how profoundly these network linkages influence the ultimate SEC forecast. The analysis reveals that the CNN's three separate channels independently extract highly distinct insights from the raw data. Specifically, the initial channel is tuned to inputs arriving at 1- and 10-minute intervals; this pathway accounts for 53.44% of the aggregate input weight and drives roughly 70.32% of the final SEC output. Meanwhile, the middle channel is responsible for detecting trends within 5- and 10-minute intervals, commanding 65.58% of the overall input weight. The terminal channel of the TLCNN is almost exclusively dedicated to the fastest, 1-minute sampling intervals. Paradoxically, despite processing the most rapidly updating data streams, network analysis proves that this final channel's direct contribution influences a mere 28.41% of the absolute SEC values.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press



© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Figure 14. Network analysis result of TLCNN.

Energy saving analysis

Any predictive framework lacking precise tracking accuracy fails when tasked with mapping true SEC behaviors. To quantify prediction error, a residual is computed by subtracting the forecasted SEC from the measured SEC, treating the model's output as the baseline expectation. A ballooning positive residual essentially signals an undesirable operational state where energy utilization spikes while actual production lags. When mapping the actual versus forecasted SEC generated by our advanced model, significant energy gaps (residuals) become plainly visible. As depicted in **Figure 15**, distinct anomalies emerge where the residual drastically breaches the 95th percentile threshold relative to proximate data points. These anomalous spikes can be systematically partitioned into two distinct groupings, dictated by their respective ethylene plant loads (cluster 1 is 55,000 kg/h and cluster 2 is 52,300 kg/h).

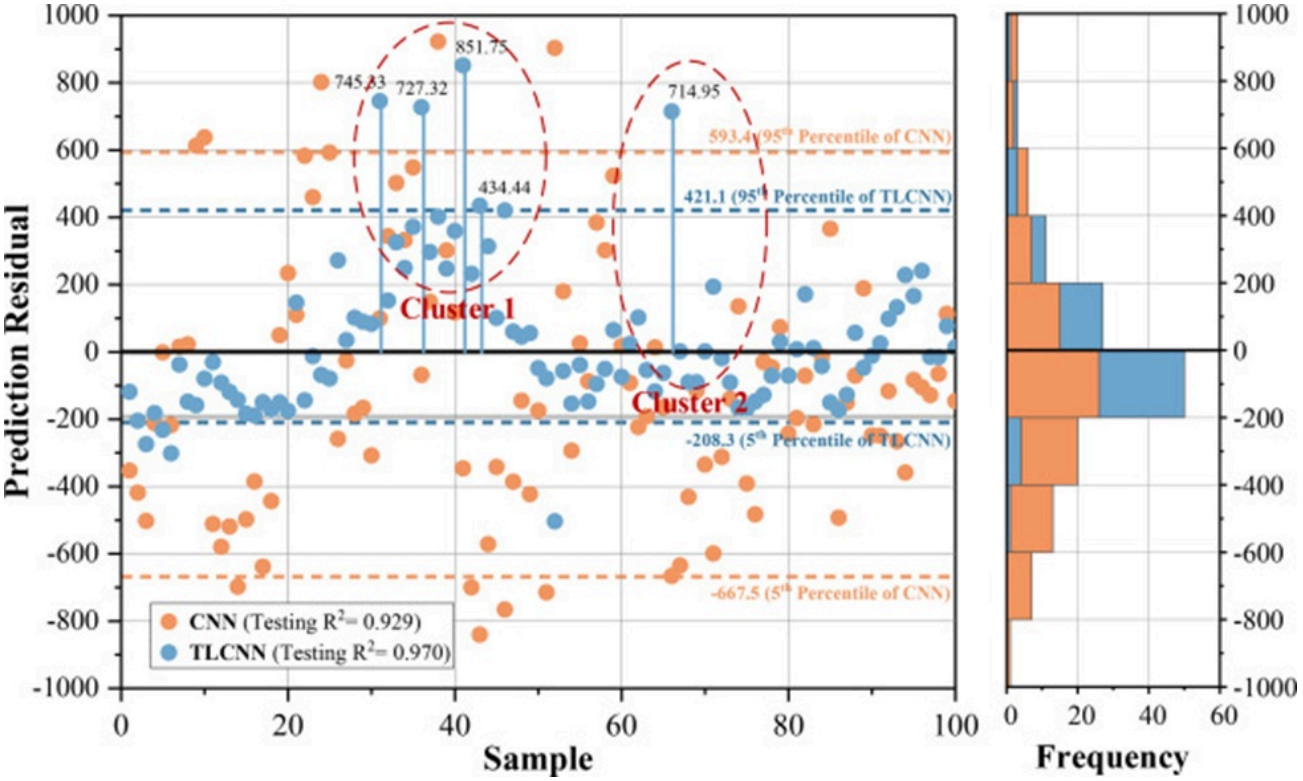


Figure 15. Residual plot of actual and benchmark SEC provided by models.

Focusing on the critical variables cataloged in **Table 4**, **Figure 16a** maps out the optimized, energy-efficient operational bounds for cluster 1. For the primary outlier in this group (the 31st sample), the recorded SEC hits 21,413 kJ/kg, wildly eclipsing the expected benchmark of roughly 20,668 kJ/kg. When this 31st sample is evaluated against neighboring, efficiently performing points (those residing under the 95th percentile boundary with identical ethylene loads), it becomes glaringly obvious that its energy draw is abnormally high compared to the established solution space (indicated by the shaded region). Within this optimal zone, normal SEC spans a tight window of 20,184 to 20,764 kJ/kg. By deliberately manipulating and constraining the process parameters to match this space, the facility could slash its SEC by an impressive 649 to 1,229 kJ/kg. Executing this same analytical protocol on cluster 2 (the 43rd sample) shown in **Figure 16b** yields similar insights: the 43rd sample operates at an inflated SEC of 24,063 MJ/ton, while the ideal solution space hovers safely between 22,611 and 23,654 kJ/kg. This uncovers a tangible conservation potential of 409 to 1,451 kJ/kg, representing a massive 3.5% to 11% reduction in annual utility expenditures. Projecting these findings into macro-level production planning and process management unlocks an estimated yearly financial savings of 299,614 USD—calculated directly using natural gas spot market pricing provided by the U.S. Energy Information Administration [43]. Environmentally, averting this excess energy burn stops 5709.59 tons of carbon dioxide equivalent from entering the

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

atmosphere annually, a figure derived using official organizational greenhouse gas reporting multipliers from the U.S. Environmental Protection Agency [44]. Based on these standards, the specific emission multipliers applied for CO2 (1 GWP), CH4 (25 GWP), and N2O (298 GWP) correspond to 50,291.18, 947.82, and 94.78 kg/TJ, respectively.

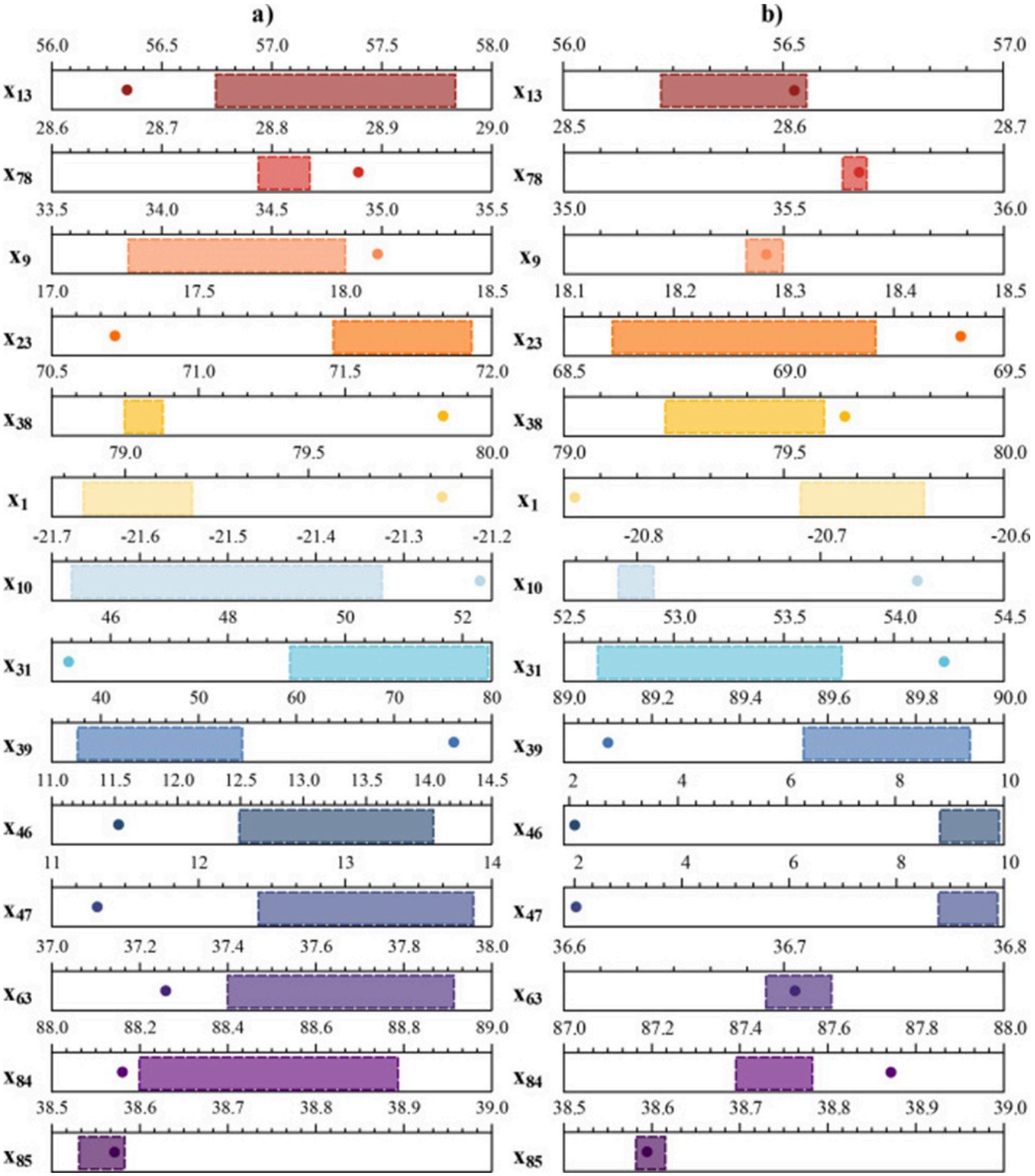


Figure 16. Benchmarking of energy-efficient operating range for a) cluster 1 and b) cluster 2.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Conclusion

This research successfully formulated TLCNN models dedicated to projecting energy efficiency and executing energy-conservation diagnostics for the VCM process, specifically handling the complexities of multirate measurement intervals. Integrating a multi-channel architecture significantly boosted the system's analytical precision, allowing individual pathways to capture distinct data trends independently and thereby expanding overall learning capacity. By capitalizing on transfer learning, the architecture dramatically improved its agility when processing foreign datasets, gracefully mitigated post-training timeframe disruptions, and bypassed historical weight initialization roadblocks. Collectively, these architectural decisions vastly reinforced the model's operational reproducibility. The system's robustness was rigorously audited through real-time deployment, sliding scale timeframe checks, and replicability testing. Ultimately, synthesizing the algorithmic forecasts with exploratory data diagnostics yielded a comprehensive energy-conservation assessment, driving the following core deductions:

1. The TLCNN framework accurately digests multirate data streams to monitor the ongoing SEC of the VCM facility. While the TLCNN architecture requires 1.25 times longer training than its 3C CNN counterpart, it compensates by producing predictions 1.15 times faster and improving overall forecasting precision by 19%. This advanced model eclipses legacy learning approaches, registering superior performance metrics of 0.97 for R2, 232.862 for RMSE, and 0.784 for MAPE.

2. The specific training protocol utilized for the TLCNN delivers exceptional reproducibility and acts as a highly dependable countermeasure against weight initialization failures. By exporting all bias parameters and weighting factors via a model parameter-driven transfer learning mechanism, the TLCNN reliably locks in testing accuracies bounded between 0.929 and 0.992 without triggering a single training collapse.

3. Post-training, the proposed system is perfectly capable of weathering shifts in sampling timeframes, provided the overarching combinations of predictive variables mirror the original training matrix. Crucially, when faced with novel data combinations, the TLCNN defends its predictive accuracy by freezing the pre-trained segments of the network and forcing the supplementary fine-tuning layers to master the unfamiliar information patterns independently.

4. Executing an energy gap diagnostic—which juxtaposes the true SEC against the benchmark SEC modeled by the TLCNN—exposes a massive 3.5% to 11% margin for energy conservation. For facility leadership, capitalizing on this margin translates to roughly 300,000 USD in slashed operational costs for production management and an environmental victory of eliminating roughly 5,709 tons of CO2 equivalent emissions every year.

In summation, the engineered model delivers elite predictive accuracy, unwavering reproducibility, and a layer of explainability not typically found in conventional algorithmic approaches. It readily manages forecasting duties even when raw data is collected at chaotic, inconsistent intervals; however, its robustness is challenged when entirely foreign combinations of input variables are introduced. Looking forward, the system's resilience could be significantly improved by embedding self-adaptive learning mechanisms directly into the core structure. A network of this nature would autonomously recalibrate itself the moment predictive accuracy degrades during live deployment. By embracing this continuous adjustment, the model would maintain peak effectiveness regardless of how severely the live predictor variables deviate from the original training dataset.

Acknowledgements

The authors would like to acknowledge the support of the Faculty of Engineering, Kasetsart University (Grant No.66/04/CHEM/D.Eng), Center for Advanced Studies in Industrial Technology, and the Center of Excellence on Petrochemical and Materials Technology. We would also like to acknowledge Honeywell and GRD Tech Co., Ltd for providing the UniSim Design Suite simulation software used in the study.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 03 Mar 2023 Revised: 09 May 2023 Accepted: 26 May 2023
Published online: 15 November 2024

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

References

1. Ritchie H, Roser M, Rosado P. Energy [Internet]. Our World in Data; 2022 [cited 2026 Jul 6]. Available from: https://ourworldindata.org/energy

2. Patterson M, Singh P, Cho H. The current state of the industrial energy assessment and its impacts on the manufacturing industry. Energy Rep. 2022;8:7297-311.

3. Energy Information Administration. Manufacturing Energy Consumption Survey (MECS) 2021 [Internet]. [cited 2022 Jul 8]. Available from: https://www.eia.gov/consumption/manufacturing/data/archive/.

4. Menghi R, Papetti A, Germani M, Marconi M. Energy efficiency of manufacturing systems: a review of energy assessment methods and tools. J Clean Prod. 2019;240:118276.

5. Golmohamadi H. Demand-side management in industrial sector: a review of heavy industries. Renew Sustain Energy Rev. 2022;156:111963.

6. Gao H, Li Z, Qiu S, Yang B, Li S, Wen Y. Energy exchange efficiency prediction from non-linear regression for membrane-based energy-recovery ventilator cores. Appl Therm Eng. 2021;197:117353.

7. Zhu L, Chen J. Development of energy efficiency principal component analysis model for factor extraction and efficiency evaluation in large-scale chemical processes. Int J Energy Res. 2019;43:814-28.

8. Cattani G. Combining data envelopment analysis and Random Forest for selecting optimal locations of solar PV plants. Energy AI. 2023;11:100222.

9. Yang J, Chen B. Energy efficiency evaluation of wastewater treatment plants (WWTPs) based on data envelopment analysis. Appl Energy. 2021;289:116680.

10. Vieira LW, Marques AD, Schneider PS, José da Silva Neto A, Viana FAC, Abdel-jawad M, et al. Methodology for ranking controllable parameters to enhance operation of a steam generator with a combined Artificial Neural Network and Design of Experiments approach. Energy AI. 2021;3:100040.

11. Walser T, Sauer A. Typical load profile-supported convolutional neural network for short-term load forecasting in the industrial sector. Energy AI. 2021;5:100104.

12. Machalek D, Tuttle J, Andersson K, Powell KM. Dynamic energy system modeling using hybrid physics-based and machine learning encoder–decoder models. Energy AI. 2022;9:100172.

13. Liu J, Chen L, Xu W, Feng M, Han Y, Xia T, et al. Novel production prediction model of gasoline production processes for energy saving and economic increasing based on AM-GRU integrating the UMAP algorithm. Energy. 2023;262:125536.

14. Han Y, Cao L, Geng Z, Ping W, Zuo X, Fan J, et al. Novel economy and carbon emissions prediction model of different countries or regions in the world for energy optimization using improved residual neural network. Sci Total Environ. 2023;860:160410.

15. Han Y, Du Z, Geng Z, Fan J, Wang Y. Novel long short-term memory neural network considering virtual data generation for production prediction and energy structure optimization of ethylene production processes. Chem Eng Sci. 2023;267:118372.

16. Wang S, Fan Y, Jin S, Takyi-Aninakwa P, Fernandez C. Improved anti-noise adaptive long short-term memory neural network modeling for the robust remaining useful life prediction of lithium-ion batteries. Reliab Eng Syst Saf. 2023;230:108920.

17. Wang S, Takyi-Aninakwa P, Jin S, Yu C, Fernandez C, Stroe D-I. An improved feedforward-long short-term memory modeling method for the whole-life-cycle state of charge prediction of lithium-ion batteries considering current-voltage-temperature variation. Energy. 2022;254:124224.

18. Li J, Lin M, Li Y, Wang X. Transfer learning network for nuclear power plant fault diagnosis with unlabeled data under varying operating conditions. Energy. 2022;254:124358.

19. Liu J, Hou G-Y, Shao W, Chen J. A supervised functional Bayesian inference model with transfer-learning for performance enhancement of monitoring target batches with limited data. Process Saf Environ Prot. 2023;170:670-84.

20. Mondal S, Chattopadhyay A, Mukhopadhyay A, Ray A. Transfer learning of deep neural networks for predicting thermoacoustic instabilities in combustion systems. Energy AI. 2021;5:100085.

21. Peirelinck T, Kazmi H, Mbuwir BV, Hermans C, Spiessens F, Suykens J, et al. Transfer learning in demand response: a review of algorithms for data-efficient modelling and control. Energy AI. 2022;7:100126.

22. Panjapornpon C, Bardeeniz S, Hussain MA, Chomchai P. Explainable deep transfer learning for energy efficiency prediction based on uncertainty detection and identification. Energy AI. 2023;12:100224.

23. Wang S, Ren P, Takyi-Aninakwa P, Jin S, Fernandez C. A critical review of improved deep convolutional neural network for multi-timescale state prediction of lithium-ion batteries. Energies. 2022;15:5053.

24. Heo J, Song K, Han S, Lee D-E. Multi-channel convolutional neural network for integration of meteorological and geographical features in solar power forecasting. Appl Energy. 2021;295:117083.

25. Chiney A, Paduri AR, Darapaneni N, Kulkarni S, Kadam M, Kohli I, et al. Handwritten data digitization using an anchor based multi-channel CNN (MCCNN) trained on a hybrid dataset (h-EH). Procedia Comput Sci. 2021;189:175-82.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

26. Quest H, Cauz M, Heymann F, Rod C, Perret L, Ballif C, et al. A 3D indicator for guiding AI applications in the energy sector. Energy AI. 2022;9:100167.

27. Panjapornpon C, Bardeeniz S, Hussain MA. Improving energy efficiency prediction under aberrant measurement using deep compensation networks: a case study of petrochemical process. Energy. 2022;125837.

28. Huang K, Wu S, Li Y, Yang C, Gui W. A multi-rate sampling data fusion method for fault diagnosis and its industrial applications. J Process Control. 2021;104:54-61.

29. Panjapornpon C, Bardeeniz S, Hussain MA. Deep learning approach for energy efficiency prediction with signal monitoring reliability for a vinyl chloride monomer process. Reliab Eng Syst Saf. 2023;231:109008.

30. Thebelt A, Wiebe J, Kronqvist J, Tsay C, Misener R. Maximizing information from chemical engineering data sets: applications to machine learning. Chem Eng Sci. 2022;252:117469.

31. Liu H, Han L, Cao Y. Improving transmission efficiency and reducing energy consumption with automotive continuously variable transmission: a model prediction comprehensive optimization approach. Appl Energy. 2020;274:115303.

32. Tien PW, Wei S, Darkwa J, Wood C, Calautit JK. Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality – a review. Energy AI. 2022;10:100198.

33. Moghadasi M, Ozgoli HA, Farhani F. Steam consumption prediction of a gas sweetening process with methyldiethanolamine solvent using machine learning approaches. Int J Energy Res. 2021;45:879-93.

34. Han Y, Wu H, Geng Z, Zhu Q, Gu X, Yu B. Review: energy efficiency evaluation of complex petrochemical industries. Energy. 2020;203:117893.

35. Müller IM. Feature selection for energy system modeling: identification of relevant time series information. Energy AI. 2021;4:100057.

36. Li L, Wang J. Research on feature importance evaluation of wireless signal recognition based on decision tree algorithm in cognitive computing. Cogn Syst Res. 2018;52:882-90.

37. Himeur Y, Elnour M, Fadli F, Meskin N, Petri I, Rezgui Y, et al. Next-generation energy systems for sustainable smart cities: roles of transfer learning. Sustain Cities Soc. 2022;85:104059.

38. Pinto G, Wang Z, Roy A, Hong T, Capozzoli A. Transfer learning for smart buildings: a critical review of algorithms, applications, and future perspectives. Adv Appl Energy. 2022;5:100084.

39. Chinprasit J, Panjapornpon C. Model predictive control of vinyl chloride monomer process by Aspen Plus Dynamics and MATLAB/Simulink co-simulation approach. IOP Conf Ser: Mater Sci Eng. 2020;778:012080.

40. Zhang X, Hou L, Liu J, Yang K, Chai C, Li Y, et al. Energy consumption prediction for crude oil pipelines based on integrating mechanism analysis and data mining. Energy. 2022;254:124382.

41. Lawrence A, Thollander P, Andrei M, Karlsson M. Specific Energy Consumption/Use (SEC) in energy management for improving energy efficiency in industry: meaning, usage and differences. Energies. 2019;12:247.

42. Machlev R, Heistrene L, Perl M, Levy KY, Belikov J, Mannor S, et al. Explainable Artificial Intelligence (XAI) techniques for energy and power systems: review, challenges and opportunities. Energy AI. 2022;9:100169.

43. Energy Information Administration. Henry hub natural gas spot price (Dollars per Million Btu) 2022 [Internet]. [cited 2022 Jul 8]. Available from: https://www.eia.gov/dnav/ng/hist/rngwhhdm.htm.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press

44. United States Environmental Protection Agency. GHG emission factors hub 2015 [Internet]. [cited 2022 Jul 8]. Available from: https://www.epa.gov/climateleadership/ghg-emission-factors-hub.

© 2024 The Author(s). | Published by Institute of Smart Energy Systems Press