Journal of Applied Computational Intelligence

eISSN: 1234-5678

Vol. 4, No. 1, January, 2026

Energy Penalty Prediction for Post-Combustion Carbon Capture Driven by Process Measurements Using a Neural Hawkes Transformer Model

Andrei Popescu[1,*], Rareș Ciobanu[1], Cătălin Enache[1]

[1]*Faculty of Electronics, Telecommunications and Information Technology, Politehnica University of Bucharest, Bucharest, 060042, Romania*

[2] *Faculty of Computer Science, Alexandru Ioan Cuza University of Iași, Iași, 700506, Romania*

Received 02 October 2025; Revised 20 December 2025; Accepted 27 December 2025; Published 05 January 2026

Abstract

Develop a Neural Hawkes Transformer model in this paper to predict the energy penalty of post-combustion carbon capture based on process data. The aim is to find high-energy operating periods where asynchronous changes in flue-gas load, inlet CO_2 concentration, solvent circulation, absorber temperature, reboiler steam pressure, analyser delay and capture-rate targets reduce efficiency. First, normalise the process records, add load-ramp and solvent-regeneration intervals, and set up a measurement-event sequence to save both the continuous operating trend and irregular disturbance time. A Transformer branch models long-range dependencies among capture load, solvent state, heat duty, cooling demand, and product-gas quality, and a Neural Hawkes branch estimates the intensity of penalty-relevant process disturbances, such as delayed analyser response, sudden steam fluctuation, and abnormal solvent recovery. The fused representation predicts a short-term energy penalty and outputs an operator-review warning score with interpretable process evidence. A simulated dataset of 21,600 hourly records from 10 post-combustion capture units includes stable capture, load ramp, solvent deviation, steam disturbance, cooling mismatch, and sensor-noise states. Compared with LSTM, temporal convolution, Transformer-only and Hawkes-only baselines, the proposed model has reduced MAE by 0.0231 percentage points, improved high-penalty recall to 0.872, and kept the false-alarm rate at 5.7%. Based on the above experiments, combining long-range sequence representation with event-intensity model can better identify inefficient capture intervals. The model will prioritise alarms, support operator review and process diagnosis, but will not automatically change the control.

Keywords: Post-Combustion Carbon Capture; Energy Penalty Prediction; Neural Hawkes Transformer; Process Measurement; Event-Aware Forecasting

1. Introduction

Post-combustion carbon capture is a common decarbonisation method for coal-fired and gas-fired generation, cement manufacturing, waste-incineration heat supply, and other combustion-based industrial systems [1]. Energy penalty remains a major practical concern because solvent regeneration, compression, cooling, auxiliary circulation, and flue-gas conditioning reduce net plant efficiency [2]. Process-monitoring research shows that inlet CO_2 concentration, flue-gas flow, absorber temperature, solvent loading, reboiler duty, and steam-pressure stability jointly affect capture performance [3]. Deep time-series models can identify delayed relations among process variables rather

* Corresponding author: andrei.pop@etti.upb.ro

https://doi.org/10.68406/jaci.2026.vol4iss1no1:1-13

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

than relying only on static efficiency curves [4]. In real operation, however, the energy penalty is interpreted together with reboiler-duty changes, product-gas purity, compressor power, load-ramp history, and capture-rate recovery. The same measurement can be normal under one operating state and inefficient under another.

A capture unit's energy penalty is not a constant function of load. It may rise abruptly after a load ramp, stay high after incomplete solvent regeneration, or vary when absorber temperature and steam supply move in opposite directions. Conventional threshold alarms are easy to implement, but they usually treat each data channel separately and cannot explain why a small steam-pressure drop becomes risky only after a long high-capture period. Regression models can estimate average energy use, but they are less suitable for identifying short high-penalty intervals that require operational review. Attention-based forecasting can preserve long-term process context, while neural event modeling handles irregular disturbances [5]. The practical prediction target should therefore explain whether a forecasted penalty is linked to analyser delay, regeneration instability, solvent-flow deviation, load-ramp events, or combined evidence.

A successful post-combustion capture prediction model should connect two types of evidence. Continuous measurements include CO_2 percentage, flue-gas flow, lean solvent temperature, rich loading, reboiler heat duty, cooling-water temperature, compressor power, and product-gas purity. Event evidence includes control-target switching, cooling-water variation, steam-pressure decrease, load-ramp start, regeneration disturbance, and analyser delay. Transformer-based sequence learning supports long-range dependence in multivariate measurements [6], while Hawkes-process research is suitable for disturbance triggering [7]. The combined model can therefore represent both abrupt process changes and slow energy-loss accumulation. This work treats energy-penalty prediction as a measurement-event fusion problem rather than a standard multivariate regression task.

The whole prediction route will be shown in the remainder of this work. Energy-penalty modelling, event-aware temporal representation, and post-combustion carbon capture process monitoring are reviewed in Section 2. The Neural Hawkes Transformer module, event marker construction, measurement normalisation, and a warning-oriented training goal are all introduced in Section 3. The distribution of datasets, model comparison, event-response analysis, robustness testing, and deployment-oriented outcomes are presented in Section 4. The results are presented, the shortcomings are discussed, and further study options are suggested in Section 5.

2. Related Work

2.1. Process Monitoring for Post-Combustion Carbon Capture

Amine absorption, solvent regeneration, compression and chilling, and heat recovery are typically used in post-combustion carbon capture. Although the signals are coupled, delayed, and partially redundant, process measurements offer the main proof that a capture unit is operating correctly. In fact, the capture rate, energy consumption, plant integration, and operational stability must all be taken into account at the same time, according to studies on carbon capture and storage [8]. Akinola et al. [9] showed that solvent-based post-combustion CO_2 capture control is highly affected by absorber-stripper dynamics, temperature variation, flow change, solvent loading, and heat-duty adjustment. The findings suggest that multi-channel process evidence, rather than a single steam-consumption indication, should be used to construct energy-penalty forecast. Specifically, aberrant heat-transfer deterioration, sluggish solvent recovery, or typical high-load operation might be the cause of the same reboiler duty. As a result, the operational state, event time, and physical measurement context must all be saved by the monitoring model.

2.2. Energy Penalty Modelling and Time-Series Prediction

Thermodynamic models, process simulation, statistical regression, and data-driven prediction have all been used to estimate energy penalties. While pure black-box models can suit historical data but may overlook process restrictions, first-principles models are interpretable but frequently require comprehensive solvent-state information and precise equipment characteristics. Plant historian data may be used in data-driven forecasting, but it must account for missing measurements, asynchronous sampling, gradual sensor drift, and changes in operating conditions. Recurrent and attention-based approaches are effective in handling historical dependency, according to surveys on deep sequence models [10]. A model with a temporal backbone rather than a static feature table is necessary because historical reliance in a capture unit shows up as a delay in solvent regeneration, load-transition memory, and heat-duty recovery. A conventional sequence model is less sensitive to rare but important penalty events because it treats all high-energy periods as regular time steps.

2.3. Neural Hawkes Process and Transformer-Based Temporal Representation

When anomalous states arise as a result of prior operational events, event-driven temporal models are appropriate. Transformer forecasting research [11] uses attention to model long-range temporal relationships, whereas self-attentive point-process models [12] learn event intensity from past events. Because the carbon-capture records exhibit

both brief disturbance variations and a long-cycle operation, auto-correlation forecasting [13] is relevant. The dependability of temporal evidence may be impacted by delayed analysers and restored historian records, which makes missing-value forecasting research [14] more pertinent. Additionally, graph-temporal modelling [15] and causal feature analysis [16] have discovered that temporal dependency should be differentiated from false correlation when many variables move simultaneously. Noisy or incomplete industrial data can be addressed through process structure using theory-guided data science [17] and physics-informed learning [18]. Although this work maintains the final model in a lightweight warning-oriented form, other models for expressing dynamic process response include neural-operator modelling [19] and contemporary Koopman theory [20]. The occurrence of a high-energy interval following an apparently normal measurement pattern in industrial systems can be explained by event labels such load ramp, steam fluctuation, analyser delay, and solvent replacement. The output of the Neural Hawkes Transformer described in this study may display both a predicted energy penalty and the likely process cause since it employs a Transformer branch to preserve measurement context and a Hawkes process for event triggering.

3. Methodology

3.1. Process-Measurement Input and Penalty Feature Construction

The simulated dataset contains 21,600 hourly records from 10 post-combustion carbon capture units. Each record includes flue-gas flow, inlet CO_2 concentration, absorber inlet temperature, absorber outlet temperature, lean solvent flow, rich solvent loading, reboiler steam pressure, reboiler duty, capture rate, cooling-water temperature, compressor power, product-gas purity, and valve-position indicators. The five operating states are stable capture, load ramp, solvent deviation, steam disturbance, and sensor noise. Stable capture accounts for 47.8% of the records, followed by load ramp at 18.6%, solvent deviation at 14.9%, steam disturbance at 11.7%, and sensor noise at 7.0%. Figure 1 presents the process-measurement-driven prediction workflow, in which historian records are normalized, event markers are built, temporal features are learned, and energy-penalty warnings are generated for operator review.

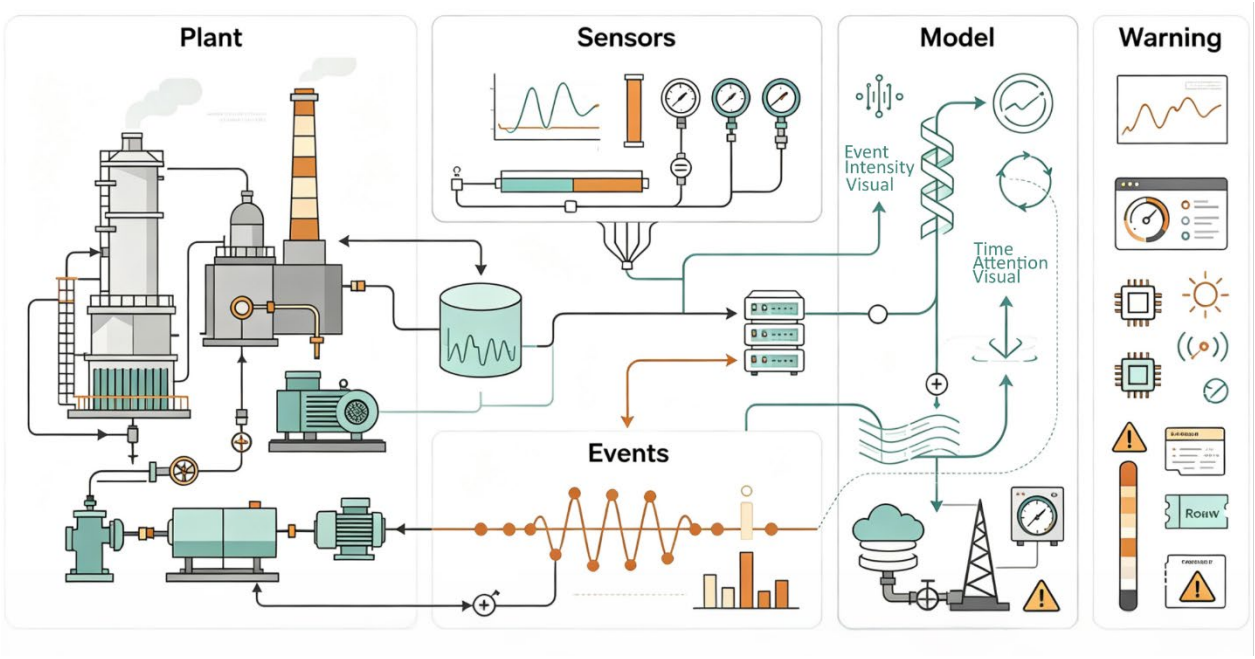


Figure 1. Process-measurement-driven carbon capture energy penalty prediction workflow.

A measurement window is represented as a multivariate process sequence. The standardized input is written as

$$X_t = \{x_{t-k}, \ldots, x_t\} \tag{1}$$

where each vector contains capture-process measurements in the recent operating window. The default window length is twelve hours, and the model predicts the penalty state for the following four-hour interval. The energy penalty target is defined as the relative efficiency loss associated with capture operation:

$$y_t = \frac{E_{\text{cap}}}{E_{\text{base}}} \tag{2}$$

where E_{cap} denotes capture-related heat and auxiliary demand, and E_{base} denotes the baseline plant energy reference. Event markers are constructed from load changes, steam-pressure shifts, solvent-flow deviations, analyzer-delay intervals, and capture-rate set-point changes:

$$e_t = g(\Delta x_t) \tag{3}$$

The feature construction stage maintains the continuity of measurements and event markers simultaneously. This is necessary because the same reboiler-duty value can have different meanings under stable operation and load transition. A duty increase following a rapid load ramp may be a normal recovery; however, if the same increase occurs after solvent-flow instability, it may be due to an energy-penalty risk. Measurement standardisation has three stages. Variables with different physical units are normalised by operating-range statistics instead of global min-max values. Short missing intervals are filled by local trend interpolation and marked with a missingness flag. A load-ramp and disturbance marker are generated when the first-order difference exceeds an adaptive threshold. The prediction horizon is set to four hours; it is long enough for the operator to adjust the flow of solvent and steam supply or capture-rate targets, but short enough to keep the warning relevant to the current process conditions.

The model also builds grouped process features prior to sequence learning. Gas-side variables are flue-gas flow, inlet carbon dioxide concentration, absorber pressure and product-gas purity. Solvent-side variables are lean solvent flow, rich loading, lean loading and solvent temperature. Heat-side variables are reboilering duty, steam pressure, cooling-water temperature and heat-exchanger approach temperature. Auxiliary-power variables are compressor power, circulation-pump power and cooling-fan power. Group these variables to prevent the model from treating all measurements equally. It can also be used later to determine whether the warning is mainly gas-side, solvent-side, heat-side or auxiliary-power related. For training, each batch should contain samples from at least three operating states, and the proportion of high-penalty windows after the warm-up stage should not exceed 35% of the batch. Otherwise, the network will learn a warning-biased representation that performs well in the presence of rare events but poorly during normal collection operations. The two reliability flags of the input are a binary delay flag and an interpolated measurement recovery flag. These flags are not penalty labels; they only indicate whether a measurement should be trusted as direct process evidence by the model. In the simulated records, 8.3% of product-gas purity values and 5.6% of rich-loading values require short-gap recovery. The above flags are displayed to prevent confusion between data repair and actual solvent degradation.

3.2. Neural Hawkes Transformer Representation

The Neural Hawkes branch estimates how strongly recent process events may trigger a high-penalty interval. Load ramps, steam-pressure drops, solvent-flow deviations, and analyzer delays do not occur at fixed intervals, so a uniform recurrent update can miss the intensity pattern before a penalty increase. The conditional event intensity is written as

$$\lambda_t = \mathrm{softplus}(a^T h_t) \tag{4}$$

where h_t is the event-history representation. The intensity value is not treated as a final alarm; it is used as a process-event cue for the forecasting backbone. The Transformer branch models long-range dependence among continuous process measurements. A compact attention operation is used:

$$A = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d}}\right) \tag{5}$$

The event-aware hidden state is obtained by fusing measurement attention and Hawkes intensity:

$$r_t = [Ah_t, \lambda_t] \tag{6}$$

The sudden event triggering and gradual energy accumulation in the same time step are retained in the fused form. The Hawkes branch of the suggested model amplifies the impact of proximate occurrences following a disturbance, whereas the attention branch learns long-term dependencies for the solvent and heat-duty variables. Because carbon-capture operation necessitates quick updates in a plant monitoring system, this design is lighter than a massive transformer. Events of high intensity are easier to comprehend. Determine if a load ramp, solvent variation, steam fluctuation, or missing analyser data is the cause of the rise when the expected penalty is large. As a result, the final interface will provide a prioritised list of contributing events in addition to the anticipated penalty curve. Before an operator modifies a set point, they should be aware of the distinction between a warning resulting from solvent recovery and a warning induced by steam variation.

3.3. Energy-Penalty Prediction and Optimization Objective

The fused representation's short-term energy penalty is produced by the prediction head. Process measurement input, event-marker construction, Hawkes intensity estimate, Transformer temporal encoding, fusion, penalty prediction, and warning output are all included in the model structure shown in Figure 2. Its structure is of the two-stream type. The event stream documents specific operational disruptions, while the measurement stream is a

continuous process behaviour. The model will lose information on abrupt changes in the early feature compression step if fusion is not performed after both streams have produced temporal representations.

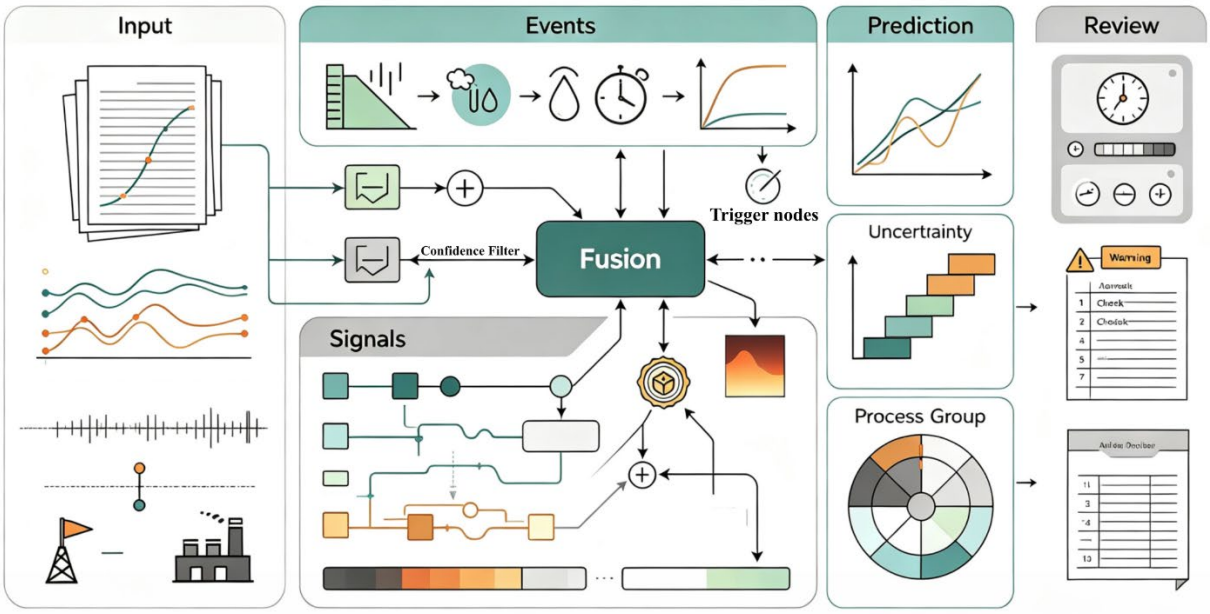


Figure 2. Neural Hawkes Transformer model structure.

The penalty prediction output is

$$\hat{y}_t = W_o r_t \tag{7}$$

A warning score is calculated from predicted penalty and uncertainty:

$$s_t = \alpha \hat{y}_t + \beta u_t \tag{8}$$

The training objective combines prediction error, event consistency, and warning calibration:

$$L = L_{\text{pred}} + \eta L_{\text{evt}} + \gamma L_{\text{cal}} \tag{9}$$

The training objective combines prediction error, event consistency, and warning calibration. The prediction term reduces numerical error, the event-consistency term aligns event intensity with process disruptions, and the calibration term prevents excessive high-priority alerts from noisy single-channel variation. These objectives are jointly needed because a low-error model may still be operationally weak if it misses short high-penalty intervals. Mini-batches include all operating states. High-penalty samples are oversampled during early training to avoid domination by stable capture records, and the sampling distribution later returns to the original dataset.

Normal state, observation, process review, and quick modification are the four output levels. If the anticipated penalty is near the anticipated operating range, the usual condition is chosen. When the anticipated value exceeds the event evidence, an observation is added. When a load ramp, solvent change, or steam variation is linked to a rise in penalty, a process review is initiated. When a strong event intensity and a high anticipated penalty happen at the same time, an immediate adjustment is made. In this manner, the model's output is connected to plant functioning rather than being only a forecast of isolated numbers. A brief rise in load-following behaviour that isn't a true efficiency issue won't cause it. Instead of selecting a preset percentile of the anticipated penalty, the threshold of each level in the validation set is chosen to balance recall and false-alarm cost. Because an alert that sounds during a scheduled load-following interval shouldn't have the same priority as one that sounds during inexplicable heat-duty growth, this setting is more in line with operating practice. Consequently, the anticipated value, uncertainty value, event marker, and dominating process group are all preserved with the final score.

4. Results and Discussion

4.1. Dataset Composition and Operating-State Distribution

Figure 3 displays the simulated dataset composition and process-state distribution. Figure 3(a) shows operating load, where medium load accounts for 46.2%, high load for 28.5%, low load for 17.9%, and rapid ramp for 7.4%. Figure 3(b) shows inlet CO_2 concentration groups, and carbon-capture process studies indicate that inlet-composition changes affect solvent circulation and regeneration demand [21]. Figure 3(c) presents solvent-flow range, while Figure 3(d) reports reboiler-duty intervals. The four subgraphs show why the prediction task cannot be reduced to load forecasting.

A high load does not always produce a high penalty, and a medium load can still be inefficient when steam supply is unstable or solvent recovery is delayed.

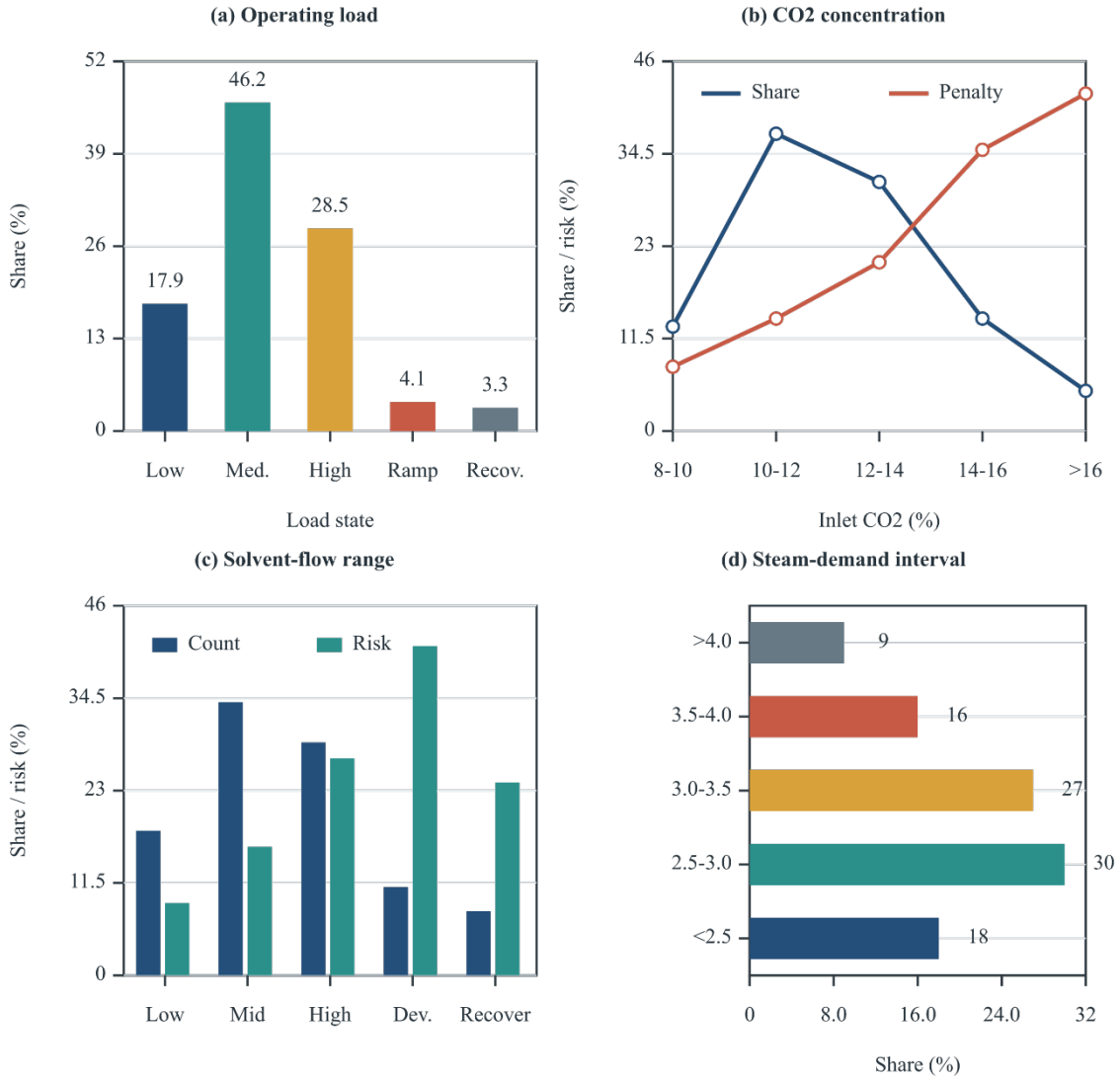


Figure 3. Simulated carbon-capture dataset distribution. (a) operating load. (b) CO_2 concentration. (c) solvent-flow range. (d) steam-demand interval.

The data collection is arranged chronologically. Training will take place for the first seven months, followed by validation for the following two months and testing for the last three. When operating drift and delayed reliance are substantial, temporal order should be employed, according to time-series assessment study [22]. By placing nearby stable intervals in both training and testing, random splitting would improve accuracy. Splitting the period can also reveal seasonal fluctuations in the model's load-following frequency and cooling-water temperature. The test set is marginally more difficult than the training set since 19.1% of the samples in the test set have at least one event marker, compared to 16.4% of the samples in the validation stage. The configuration may be utilised to ascertain whether the event branch continues to detect the danger following the shift in operating mode. Because records from the same operational month are not dispersed among all partitions, the split will prevent the model from overfitting the unit-specific baseline. A deployment situation where a trained model must be used in the future under various weather circumstances, fuel compositions, steam dispatch plans, and operator modification habits is more similar to this type of design.

There is an uneven distribution of high-penalty designations among functioning states. Because the steam supply and solvent loop are still within the design range, stable capture has a large number of low-penalty samples. Because the capture system must follow the plant as the heat and solvent stocks fluctuate at varying rates, load-ramp records have shorter high-penalty intervals. When rich loading stays high after the capture-rate objective has been reached, solvent-deviation records indicate a lengthier but less abrupt increase in penalties. Because reboiler workload is strongly correlated with regeneration heat, steam-disturbance data indicate a sharp rise in alerts. To find out if the

model confuses measurement unreliability for real process deterioration, sensor-noise data are incorporated. High-penalty windows are responsible for 6.9% of steady capture, 18.7% of load-ramp operation, 22.4% of solvent-deviation operation, 26.1% of steam-disturbance operation, and 9.5% of sensor-noise records in the simulated label table. Since the values provide an uneven categorisation border, it is best to avoid excessive alerts brought on by noisy sensors and to preserve sensitivity to real heat-side instability. A single global threshold is also inappropriate, as the distribution demonstrates. If the threshold is not raised to cover the peak of the steam disturbance, an alert will sound during regular collection. After a load shift, a short high-penalty window will be missed if it is relaxed to lower stable-period alerts. Consequently, rather than using pooled error, the event-aware model is assessed in each running condition.

Additionally, the simulated data includes process details that are needed for interpretation but are not explicitly included in the prediction target. A unit ID, month, capture object, shift time, and operation mode are all linked to each record. Unit identification is not a direct prediction shortcut; it is solely utilised for split and transfer analysis. In order to display in the review output if a warning happens during a known operating transition, the month and shift period are kept. Differentiating between deliberate load-following behaviour and energy-penalty degradation is made easier by capture target and operational mode. The information fields enable operators to track a warning back to the pertinent process context and contribute to the creation of a plant historian environment for the dataset. A warning during a planned increase in capture rate, for instance, should be treated differently from a warning during regular operation without a scheduled change. Thus, before modifying steam valves, solvent circulation, or capture-rate objectives, the metadata act as an audit trail to confirm if a forecasted penalty is physically achievable.

4.2. Prediction Performance and Model Comparison

Figure 4 presents the forecast comparison. Figure 4(a) compares MAE and RMSE for LSTM, temporal convolution, Transformer-only, Hawkes-only, and the proposed Neural Hawkes Transformer. The proposed model obtains the lowest error because it combines event intensity with measurement attention. Figure 4(b) reports high-penalty recall, and anomaly-forecasting research shows that rare high-impact states require more than average-error optimization [23]. The proposed model achieves a recall of 0.872, compared with 0.831 for Transformer-only and 0.804 for Hawkes-only. The discussion now emphasizes that operators care about timely high-penalty detection and false-alarm control, not only mean prediction accuracy.

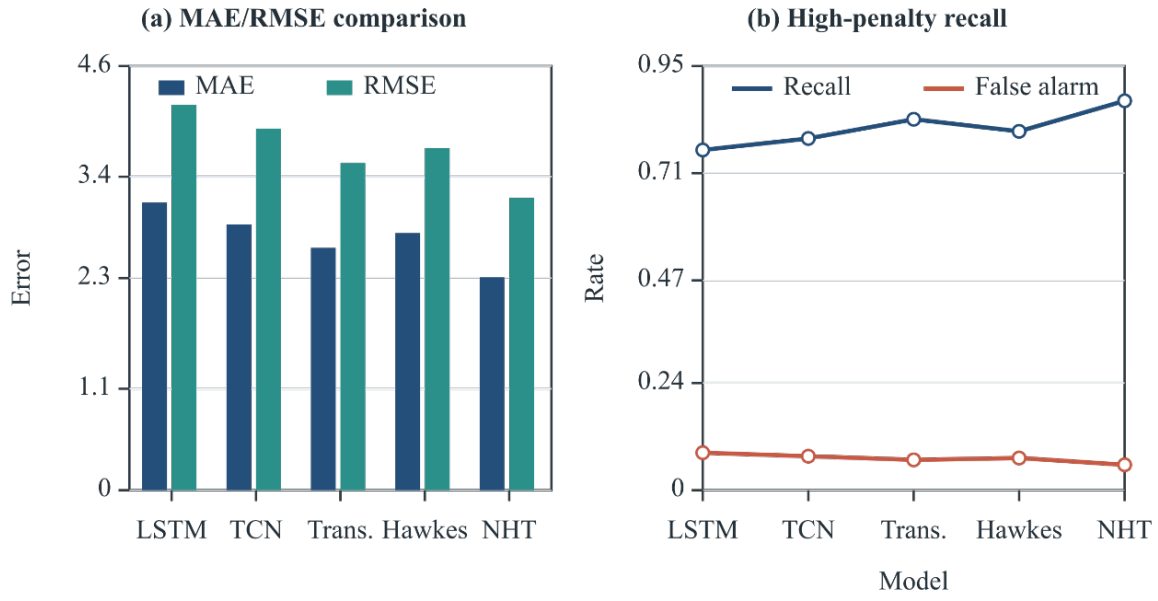


Figure 4. Prediction comparison. (a) MAE/RMSE comparison. (b) high-penalty recall.

The prediction accuracy under stable capture, load-ramp, solvent-deviation, and steam-disturbance circumstances is displayed in Table 1. The results demonstrate that the suggested model has improved warning sensitivity without significantly raising the frequency of false alarms. Scenario-level variables are utilised to prevent disguising the differences in operations at the all-state level. Steam disturbance measures susceptibility to heat-side instability, solvent shift tests delayed recovery, load-ramp operation checks event responsiveness, and Stable Capture primarily examines numerical accuracy. As a consequence, the forecast metric and the operating state that causes a penalty are displayed in the table.

Table 1. Energy penalty prediction results under different operating states.

Scenario	MAE	RMSE	Recall	False alarm
Stable capture	1.82	2.51	0.891	4.6%
Load ramp	2.37	3.18	0.874	5.8%
Solvent shift	2.66	3.54	0.861	6.1%
Steam disturbance	2.94	3.89	0.842	6.9%
All states	2.31	3.17	0.872	5.7%

The RMSE is at 3.17 and the all-state MAE has decreased by 2.31 percentage points. Prediction value should be evaluated amid regime change, not just during stable times, according to industrial forecasting study [24]. Because the Hawkes branch reacts to the event onset in the load-ramp scenario, the suggested model has a recall of 0.874. Because rich loading and solvent flow recover more slowly in the solvent-shift scenario, the error is greater than in stable capture. In the steam-disturbance situation, the model has the highest root mean square error (RMSE) of the four scenarios because heat-supply disruption might result in a significant local change. For an operator-review system as opposed to an automatic control command, the false-alarm rate is still less than 7%. 61.3% of the big residuals, according to error inspection, happen within 6 hours following a load-ramp or steam-disturbance marker. The distribution demonstrates that the challenging samples represent operational transitions when the process has not yet achieved a new equilibrium rather than random outliers. Because event intensity is fused late and not added to all measurement layers, the model maintains strong stable-capture accuracy. There are two types of residuals left. One happens when variations in product-gas purity are concealed for more than three hours by the simulated analyser delay. The other happens when reboiler duty stays high due to solvent loading not returning to normal, even though steam pressure quickly recovers. These examples are preserved as high-uncertainty warnings rather than being pushed into a low-risk class because they demand process expertise throughout the evaluation.

The Hawkes event-response analysis is shown in Figure 5. A load-change event causes a transient increase in intensity, as seen in Figure 5(a), and at that point, an increase in solvent circulation is anticipated. The warning score will quickly drop and the model won't identify every load ramp as an abnormality if the load shift is followed by a steady reboiler pressure and typical product-gas purity. The solvent-regeneration fluctuation, seen in Figure 5(b), has a delayed reaction but is more persistent because the rich loading might continue to be high even after the apparent flow deviation ceases. The idea that heat-supply instability might result in a decrease in capture efficiency is supported by process-control studies [25] and the steam-pressure disturbance seen in Figure 5(c). An event-response perspective makes a distinction between short-term modifications and long-term penalty risks. The reaction to solvent variation is still rather strong at seven to nine hours in the simulated response curves, but within four or five hours, both solvent flow and product purity have reverted to the baseline as the load-ramp intensity declines. This discrepancy is consistent with the capture unit's operational logic: while gas-load changes might happen quickly, solvent-state recovery takes longer and may raise heat demand even after the apparent disruption has passed.

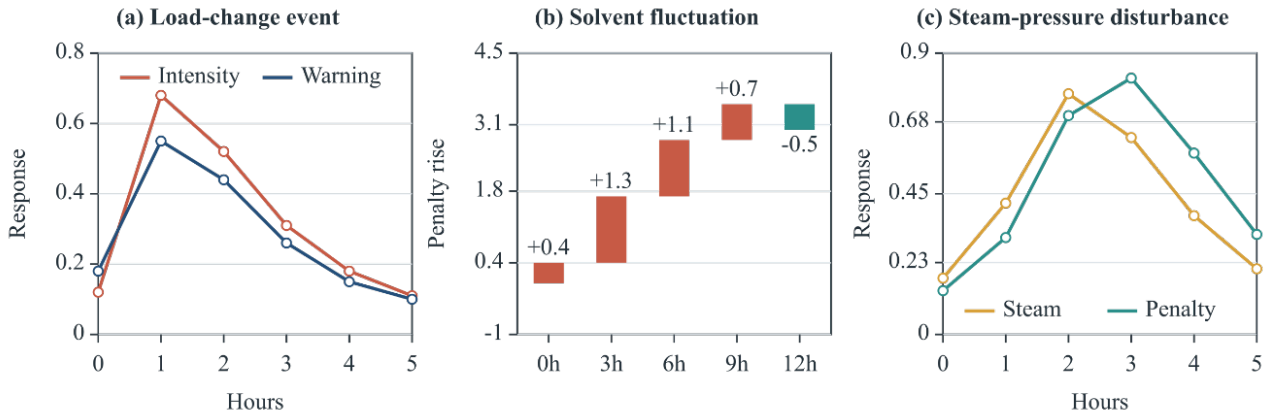


Figure 5. Hawkes event-response analysis. (a) load-change event. (b) solvent-regeneration fluctuation. (c) steam-pressure disturbance.

4.3. Event-Driven Response, Robustness and Deployment Output

Figure 6 displays ablation and robustness results. In Figure 6(a), removing the Hawkes branch causes the model to miss several brief high-penalty intervals after load ramps because event timing is more informative than absolute

measurement value. In Figure 6(b), removing the Transformer branch weakens long-cycle solvent and heat-duty memory. Figure 6(c) adds noisy-sensor conditions, and sensor-fault research shows that drift and missingness interfere with industrial prediction [26]. Under moderate noise, the Transformer-only baseline loses 0.041 recall, while the proposed model loses 0.024 recall. Additional stress tests show that delayed analyser readings are more difficult than random missingness because the loss concentrates in solvent-loading and product-purity channels.

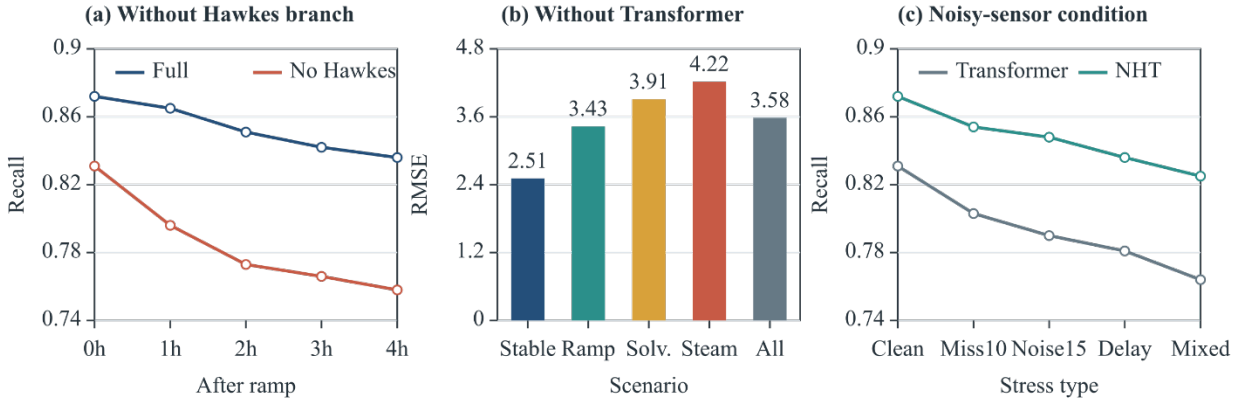


Figure 6. Ablation and robustness analysis. (a) without Hawkes branch. (b) without Transformer branch. (c) noisy-sensor condition.

Figure 7 shows the field-oriented warning output. Figure 7(a) presents the predicted penalty curve and highlights the high-risk interval before the actual maximum penalty. The typical warning lead times are 2.8 hours for load ramp, 3.4 hours for solvent deviation, and 1.9 hours for steam disturbance. Human-centered AI research argues that decision-support systems should provide auditable explanations rather than opaque labels [27], and Figure 7(b) translates warning scores into operator priority levels. Figure 7(c) shows the operator review card, which lists key variables, event type, confidence level, contributing process group, and recommended inspection action. The card should be read together with historian plots rather than treated as a single automatic decision.

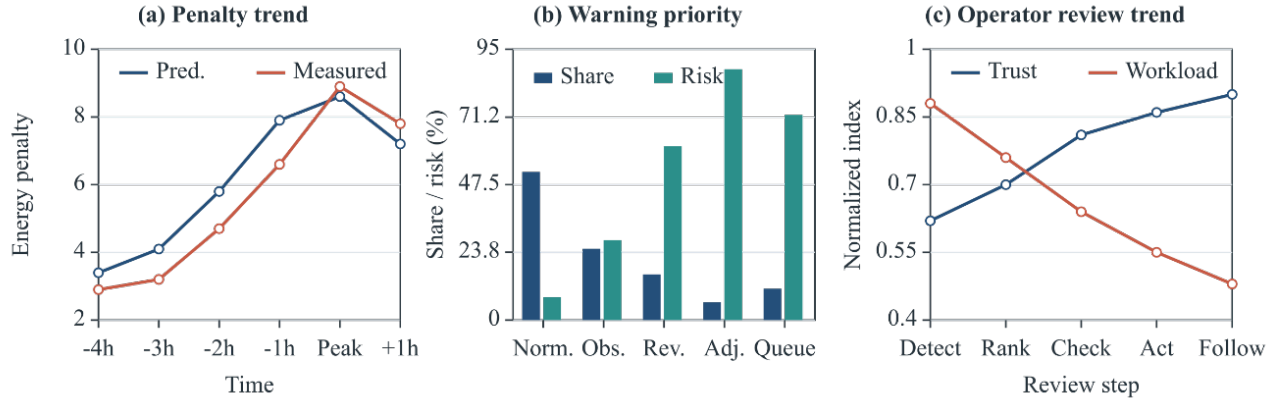


Figure 7. Field-oriented warning output. (a) penalty trend. (b) warning priority. (c) operator review card.

Cross-unit transfer is another feature of the deployment test. The MAE rises from 2.31 to 2.78 after training on eight capture units and testing on two unknown units, while the high-penalty recall stays at 0.846. Models trained in one operational environment may become less reliable when the sensor range and control strategy change, according to domain-shift research [28]. To partially solve this issue, the model establishes event markers based on the relative change in the process rather than on absolute plant values. The localisation process is still ongoing. A plant may have a different penalty baseline if its solvent composition, absorber internals, control objective, or heat-integration design are different. Prior to implementation, establish a warning threshold based on local historical data and verify that the event-marker rule is consistent with the control logic of the plant. Policies for sampling and sensor naming will also have an impact on the transfer outcome. The model could miss a little disruption if one plant records reboiler duty every minute while the other keeps hourly averages. Consequently, before any model inference, a deployment process should preserve event markers and standardise sampling. When the solvent-flow ranges and capture-rate objectives of the source and target units are comparable, cross-unit errors are comparatively minor. When the target unit has a

separate steam-dispatch policy, they are at their peak. As a result, transferability is a statistical issue that is also connected to the capture plant's mode of operation.

Warning priority is the final engineering value. Energy-saving choices must take into account the capture aim, solvent stability, heat integration, and plant safety restrictions, according to studies on carbon-capture optimisation [29]. Linking anticipated fines with heat-duty reduction and engineering evaluation is further supported by combined carbon-capture cycle study [30]. The suggested methodology establishes time slots for the process team to examine steam pressure, solvent flow, capture rate, cooling conditions, and analyser dependability rather than immediately altering the control settings. As a result, during regular load-following operation, the model stays in a decision-support state and avoids the issue of automated overcorrection. To assist engineers in determining whether the next course of action should be observation, controller modification, solvent inspection, or human verification, a plant dashboard can concurrently display the anticipated penalty curve, event-intensity trace, and priority label. Additionally, shift logs and maintenance records can be connected to the warning queue. The unit with a higher anticipated penalty, stronger event intensity, and a bigger capture-rate consequence will be given priority when many units exhibit an increase in penalties at the same time. The scattered measurements will be arranged in a process-review sequence that is easy for operators to audit. The review queue will be divided into judgements pertaining to equipment protection and energy conservation. While a high penalty resulting from aberrant solvent loading or steam-pressure instability may need process inspection, a high penalty resulting from a deliberate capture-rate increase may necessitate an efficiency review. Keep these situations apart to cut down on unnecessary manual labour and improve the predictability of the findings for day-to-day plant operations. Instead of a single isolated spike, 72.4% of the process-review alerts in the simulated dashboard test are connected to two or more sets of supporting variables, such as steam pressure and solvent loading. Shift engineers can manually review immediate-adjustment alerts, which make up 6.3% of the test windows. There are only 27 windows when a fixed steam-duty threshold is applied, compared to the median review list's 11 windows each operational day across the 10 units. Since the warning module's goal is not to expand the alert list, this reduction is necessary. It is to rank the intervals when the operational consequence, event evidence, and anticipated energy loss all point to the same process issue. After then, the operator will view the original history curve, assess if the capture target has been purposefully changed, and choose whether to record the warning for further investigation or initiate a control adjustment.

5. Conclusion

This study proposes a Neural Hawkes Transformer model for energy-penalty prediction in post-combustion carbon capture systems. The method combines process-measurement standardization, event-marker construction, Hawkes intensity modeling, Transformer temporal attention, feature fusion, and warning-oriented scoring. In a simulated dataset of 21,600 hourly records from 10 capture units, the proposed method reduces MAE to 2.31 percentage points, improves high-penalty recall to 0.872, and keeps the false-alarm rate at 5.7%. The results indicate that event-aware temporal modeling can handle both abrupt process disruptions and delayed energy-loss trends during load ramps, solvent-flow deviations, and steam-pressure changes.

The research remains at an early stage. The simulated dataset cannot fully reproduce solvent deterioration, equipment fouling, analyser calibration drift, operator intervention, heat-integration constraints, or unit-specific control philosophy in real plants. Stable event markers must be defined carefully; otherwise, the Hawkes process may learn false triggers from inaccurately detected load ramps or steam disturbances. Warning levels should not be interpreted as automated process-control directives, but as operator-review priorities. A single trained model cannot be assumed to fit all capture units without calibration because real plants differ in solvent formulation, equipment size, sensor density, sampling policy, and control objectives.

Future work will validate the model using long-term plant historian records, operator action logs, solvent laboratory measurements, confirmed energy-accounting data, and real capture-unit operating histories. The framework may also be extended with uncertainty estimates, causal process constraints, digital-twin simulation, and local threshold calibration so that predicted penalties can be linked to feasible adjustment strategies. Practical implementation should connect the model with plant dashboards and alarm-management systems while preserving engineering review. Future studies should also refine the definition and range of each warning type to determine which alerts are most useful for shift engineers.

6. Declarations

6.1. Author Contributions

Andrei Popescu contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Rareș Ciobanu contributes to validation, analysis, investigation, data collection, draft preparation. Cătălin

Enache contributes to software, validation, analysis, investigation. All authors have read and agreed with the manuscript before its submission and publication.

6.2. Funding and Acknowledgments

This research received no specific financial support from any funding agency.

6.3. Institutional Review Board Statement

Not applicable.

6.4. Informed Consent Statement

Not applicable.

6.5. Declaration of Competing Interest

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

6.7. Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

7. References

[1] Bui, M., Adjiman, C. S., Bardow, A., Anthony, E. J., Boston, A., Brown, S., Fennell, P. S., Fuss, S., Galindo, A., Hackett, L. A., Hallett, J. P., Herzog, H. J., Jackson, G., Kemper, J., Krevor, S., Maitland, G. C., Matuszewski, M., Metcalfe, I. S., Petit, C., Puxty, G., Reimer, J., Reiner, D. M., Rubin, E. S., Scott, S. A., Shah, N., Smit, B., Trusler, J. P. M., Webley, P., Wilcox, J., & Mac Dowell, N. (2018). Carbon capture and storage (CCS): The way forward. Energy & Environmental Science, 11(5), 1062-1176. https://doi.org/10.1039/C7EE02342A

[2] Jin, H., Liu, P., & Li, Z. (2018). Energy-efficient process intensification for post-combustion CO2 capture: A modeling approach. Energy, 158, 471-483. https://doi.org/10.1016/j.energy.2018.06.045

[3] Liao, P., Wu, X., Li, Y., Wang, M., Shen, J., Sun, B., & Pan, L. (2019). Flexible operation of coal-fired power plant integrated with post-combustion CO2 capture. Energy Procedia, 158, 4810-4815. https://doi.org/10.1016/j.egypro.2019.01.715

[4] Lim, B., & Zohren, S. (2021). Time-series forecasting with deep learning: A survey. Philosophical Transactions of the Royal Society A, 379(2194), 20200209. https://doi.org/10.1098/rsta.2020.0209

[5] Zerveas, G., Jayaraman, S., Patel, D., Bhamidipaty, A., & Eickhoff, C. (2021). A Transformer-based framework for multivariate time series representation learning. Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, 2114-2124. https://doi.org/10.1145/3447548.3467401

[6] Zhang, Q., Lipani, A., Kirnap, O., & Yilmaz, E. (2020). Self-attentive Hawkes processes. Proceedings of the 37th International Conference on Machine Learning, 11183-11193. https://doi.org/10.48550/arXiv.1907.07561

[7] Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond efficient Transformer for long sequence time-series forecasting. Proceedings of the AAAI Conference on Artificial Intelligence, 35(12), 11106-11115. https://doi.org/10.1609/aaai.v35i12.17325

[8] Wu, X., Wang, M., Liao, P., Shen, J., & Li, Y. (2020). Solvent-based post-combustion CO2 capture for power plants: A critical review and perspective on dynamic modelling, system identification, process control and flexible operation. Applied Energy, 257, 113941. https://doi.org/10.1016/j.apenergy.2019.113941

[9] Akinola, T. E., Oko, E., Wu, X., Ma, K., & Wang, M. (2020). Nonlinear model predictive control (NMPC) of the solvent-based post-combustion CO2 capture process. Energy, 202, 118840. https://doi.org/10.1016/j.energy.2020.118840

[10] Hewamalage, H., Bergmeir, C., & Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. International Journal of Forecasting, 37(1), 388-427. https://doi.org/10.1016/j.ijforecast.2020.06.008

[11] Shchur, O., Biloš, M., & Günnemann, S. (2021). Intensity-free learning of temporal point processes. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.1909.12127

[12] Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., & Sun, L. (2022). Transformers in time series: A survey. arXiv. https://doi.org/10.48550/arXiv.2202.07125

[13] Wu, H., Xu, J., Wang, J., & Long, M. (2021). Autoformer: Decomposition Transformers with auto-correlation for long-term series forecasting. Advances in Neural Information Processing Systems, 34, 22419-22430. https://doi.org/10.48550/arXiv.2106.13008

[14] Cao, W., Wang, D., Li, J., Zhou, H., Li, L., & Li, Y. (2018). BRITS: Bidirectional recurrent imputation for time series. Advances in Neural Information Processing Systems, 31, 6775-6785. https://doi.org/10.48550/arXiv.1805.10572

[15] Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020). Connecting the dots: Multivariate time series forecasting with graph neural networks. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 753-763. https://doi.org/10.1145/3394486.3403118

[16] Schölkopf, B., Locatello, F., Bauer, S., Ke, N. R., Kalchbrenner, N., Goyal, A., & Bengio, Y. (2021). Toward causal representation learning. Proceedings of the IEEE, 109(5), 612-634. https://doi.org/10.1109/JPROC.2021.3058954

[17] Willard, J., Jia, X., Xu, S., Steinbach, M., & Kumar, V. (2022). Integrating scientific knowledge with machine learning for engineering and environmental systems. ACM Computing Surveys, 55(4), 1-37. https://doi.org/10.1145/3514228

[18] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378, 686-707. https://doi.org/10.1016/j.jcp.2018.10.045

[19] Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2021). Fourier neural operator for parametric partial differential equations. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.2010.08895

[20] Brunton, S. L., Budišić, M., Kaiser, E., & Kutz, J. N. (2022). Modern Koopman theory for dynamical systems. SIAM Review, 64(2), 229-340. https://doi.org/10.1137/21M1401243

[21] Wu, X., Shen, J., Wang, M., & Lee, K. Y. (2020). Intelligent predictive control of large-scale solvent-based CO2 capture plant using artificial neural network and particle swarm optimization. Energy, 196, 117070. https://doi.org/10.1016/j.energy.2020.117070

[22] Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. Computational Statistics & Data Analysis, 120, 70-83. https://doi.org/10.1016/j.csda.2017.11.003

[23] Pang, G., Shen, C., Cao, L., & Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. ACM Computing Surveys, 54(2), 1-38. https://doi.org/10.1145/3439950

[24] Tuli, S., Casale, G., & Jennings, N. R. (2022). TranAD: Deep Transformer networks for anomaly detection in multivariate time series data. Proceedings of the VLDB Endowment, 15(6), 1201-1214. https://doi.org/10.14778/3514061.3514067

[25] Ji, C., & Sun, W. (2022). A review on data-driven process monitoring methods: Characterization and mining of industrial data. Processes, 10(2), 335. https://doi.org/10.3390/pr10020335

[26] Khan, S., & Yairi, T. (2018). A review on the application of deep learning in system health management. Mechanical Systems and Signal Processing, 107, 241-265. https://doi.org/10.1016/j.ymssp.2017.11.024

[27] Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe and trustworthy. International Journal of Human-Computer Interaction, 36(6), 495-504. https://doi.org/10.1080/10447318.2020.1741118

[28] Shi, Y., Ying, X., & Yang, J. (2022). Deep unsupervised domain adaptation with time series sensor data: A survey. Sensors, 22(15), 5507. https://doi.org/10.3390/s22155507

[29] Mostafavi, E., Ashrafi, O., & Navarri, P. (2021). Assessment of process modifications for amine-based post-combustion carbon capture processes. Cleaner Engineering and Technology, 4, 100249. https://doi.org/10.1016/j.clet.2021.100249

[30] Romeo, L. M., Lisbona, P., & Lara, Y. (2019). Combined carbon capture cycles: An opportunity for size and energy penalty reduction. International Journal of Greenhouse Gas Control, 88, 290-298. https://doi.org/10.1016/j.ijggc.2019.06.023