ORIGINAL RESEARCH Open access

Closed-Loop Power Load Forecasting for Mine Integrated Energy Systems Using Rescheduling Cost-Aware Adaptive Asymmetric Error Learning

Nokuthula Dlamini[1*], Sipho Zulu[1]

Abstract

Building a mining-oriented integrated energy system capable of harvesting and reusing multiple resource flows represents a cornerstone tactic for propelling the mining sector toward its twin carbon abatement milestones. The safe, high-efficiency dispatch of such a system hinges fundamentally on precise load anticipation. Current scholarship, by and large, decouples the forecasting problem from its scheduling aftermath, treating the two as siloed domains; this separation habitually gives rise to steep rescheduling penalties whenever predictions miss the mark. Seeking to bridge this divide, we put forward a closed-loop load forecasting framework that weaves in both the cost of rescheduling and the asymmetry inherent in prediction errors. A data-generation and model-specification procedure is first described, mapping real load readings, projected values, and incurred rescheduling costs to capture the latent dependency structure tying forecast deviation to economic penalty. Given that shortfalls and surpluses impinge on dispatch outcomes in fundamentally different ways, the cost model is then fused with an asymmetric treatment of forecast residuals to shape the objective function of a Bi-LSTM forecaster. A self-calibrating mechanism for blending coefficients that weight asymmetric errors is further engineered to preserve predictive fidelity. The framework is trialed on the power-load forecasting task of an integrated energy installation at a Shanxi coal mine. Findings attest to the framework's capacity to dampen rescheduling outlays without sacrificing forecasting exactitude, spotlighting its latent utility for electrical load forecasting in mining-grade integrated energy systems.

Keywords Load forecasting, Mine integrated energy system, Rescheduling cost, Asymmetric errors

*Correspondence:
Nokuthula Dlamini
dlaminin@ukzn.ac.za
[1] Department of Intelligent Energy Networks, Faculty of Engineering, University of KwaZulu-Natal, Durban, South Africa

Introduction

An Integrated Energy System (IES) [1] denotes an advanced energy architecture that marshals modern digital infrastructures and inventive governance templates to amalgamate electricity, cooling, heat, natural gas, and renewable generation within a circumscribed geographic scope. Through this convergence, the arrangement pursues harmonized planning, operational fine-tuning, cross-system management, bidirectional responsiveness, and synergistic complementarity among constituent subsystems—goals that collectively elevate energy provision, bolster conversion efficiency, and spur enduring sustainability. In the energy-intensive context of mining, erecting an IES engineered to recapture and repurpose disparate byproduct streams can lift utilization rates while depressing carbon intensity, thereby directly advancing the twin carbon milestones [2].

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.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

Where a Mine IES (MIES) departs from a conventional regional IES (RIES) is in its inclusion of derived energy vectors—spent ventilation gas, methane drained from seams, and mine water influx. The tight, intricate interdependencies among these carriers markedly heighten the complexity of scheduling operations. Furthermore, the mine site is populated by numerous high-capacity loads and machinery expressly designed for underground duty. Across the production cycle, the intermittent activation and shutdown of cutting and development equipment, combined with the rhythm of alternating operational stages, induces exceptionally frequent and large-amplitude oscillations in electrical load. Compounding the challenge, the distinctive underground working environment imposes extraordinarily stringent reliability constraints on power delivery to safety-critical infrastructure—shaft insulation, main ventilation, water drainage, and gas extraction among them. It follows that the accuracy demanded of load predictions must conform to the exacting safety thresholds inherent to mine operations [3]. When forecasting departs from reality, particularly when the error is underestimation, it can set off cascades of severe safety exposure. Thus, precise load forecasting is the essential foundation upon which efficient dispatch of a mine integrated energy system must be built [4]. Investigating how to achieve high-fidelity power-load forecasting for a comprehensive mining energy system, and how to extinguish the tendency toward under-prediction systematically, carries acute research significance.

The lineage of power load forecasting extends back many decades, and its methods can be partitioned into open-loop schemes, where dispatch-side feedback is absent, and closed-loop formulations, where the forecasting process ingests the consequences of prior scheduling decisions [5]. The open-loop branch encompasses: (1) Techniques anchored in statistical time series modeling. Dubey *et al*. [6] operationalized ARIMA and SARIMA specifications to interrogate meter data and project day-ahead electricity consumption, thereby equipping plant operators with decision-support intelligence that contributed to reining in energy usage. Aprillia *et al.* [7] propose a novel statistical load forecasting (SLF) using quantile regression random forest (QRRF), probability map, and risk assessment index (RAI) to obtain the actual pictorial of the outcome risk of load demand profile. Such tools deliver speed and effectiveness when the underlying dynamics are linear. Yet, as economic formations continue their rapid metamorphosis, amplifying the stochasticity and nonlinearity of power loads, the edge of statistical methods dulls when the terrain becomes markedly nonlinear [8]. (2) Forecasters built upon modern machine learning. In the study by Kwon *et al*. [9], STLF based on a deep neural network using an LSTM layer is proposed. Kwon *et al.* [9] observe relationships among past data and then use them to model. Huang *et al*. [10] proposed models based on feature embedding and Transformer architectures, leveraging Transformer layers to uncover nonlinear temporal dependencies, thereby enhancing model performance. Although stand-alone machine-learning algorithms have matured to the point where they can credibly represent high-dimensional, convoluted data, their fundamental inductive limitations tend to imprint structured, often irreducible, predictive biases that are not readily erased through hyperparameter manipulation [11]. (3) Composite architectures in which data-centric components are chained together to attack engineering problems through data curation, feature extraction, and model identification. Illustratively, Tan *et al*. [12] proposed an IES multi-energy forecasting method based on composite correlation analysis and load participation factors (LPF), exploring dynamic coupling among loads and improving average prediction accuracy for electricity, cooling, and heating loads. The main novelty in Venkataramana *et al*. [13] is the combination of a Recurrent Neural Network (RNN) and Principal Component Analysis (PCA) *t*echniques to improve the forecasting capability of hourly load on an electric power substation. A dominant motif running through this body of work is a pursuit of raw accuracy, benchmarked through canonical statistical aggregators such as Mean Absolute Error. Absent from this viewpoint, however, is an appreciation that surplus-oriented forecast errors and deficit-oriented forecast errors catalyze sharply different corrective expenditures: over-prediction yields unneeded generation that must be stored, dumped, or idled at cost, whereas under-prediction forces last-minute electricity procurement at premium tariffs. These divergent economic consequences—radically asymmetric in nature—render the symmetric treatment of error polarity a serious conceptual blind spot.

Recent investigations have started to confront the problem of asymmetric prediction. Asymmetrically weighted loss functions have proven useful in image processing and regression-based forecasting. Rengasamy *et al*. [14] expanded the use of a specific form of weighted loss function to estimate remaining useful life (RUL) for aircraft engines, examining the influence of four asymmetric loss functions on the efficacy of different deep learning architectures, thereby markedly boosting RUL predictions. Bora *et al*. [15] scrutinized the plausibility of the Department of Agriculture's historical net

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

income and price forecasts through a flexible multivariate loss function that permits asymmetric loss and inseparable forecast errors, thus lowering expected loss. Abolghasemi and Bean [16] examined the distinct consequences of under-forecasting and over-forecasting on optimization costs, determining the most fitting loss function via numerous trials to curtail costs linked to final decisions. Ghanim *et al*. [17] proposed multiple Long Short-Term Memory (LSTM) frameworks featuring diverse asymmetric loss functions, assigning steeper penalties to under-forecasting to reduce underestimation and forestall potential blackout-related harm. They likewise employed a Density-Based Spatial Clustering of Applications with Noise (DBSCAN) anomaly detection methodology to remove outliers before generating forecasts, thereby improving prediction fidelity. The previously cited works indicate that contemporary asymmetric forecasting efforts center primarily on the loss functions of predictive models, intensifying penalties for under-forecasting to mitigate underestimation. Nonetheless, the effect of asymmetric forecasting inaccuracies on system functioning is not purely a question of weighting. Asymmetric loss functions can only partly alleviate underestimation and are incapable of encapsulating the interplay between forecasting errors and system scheduling consequences **(Figure 1).**

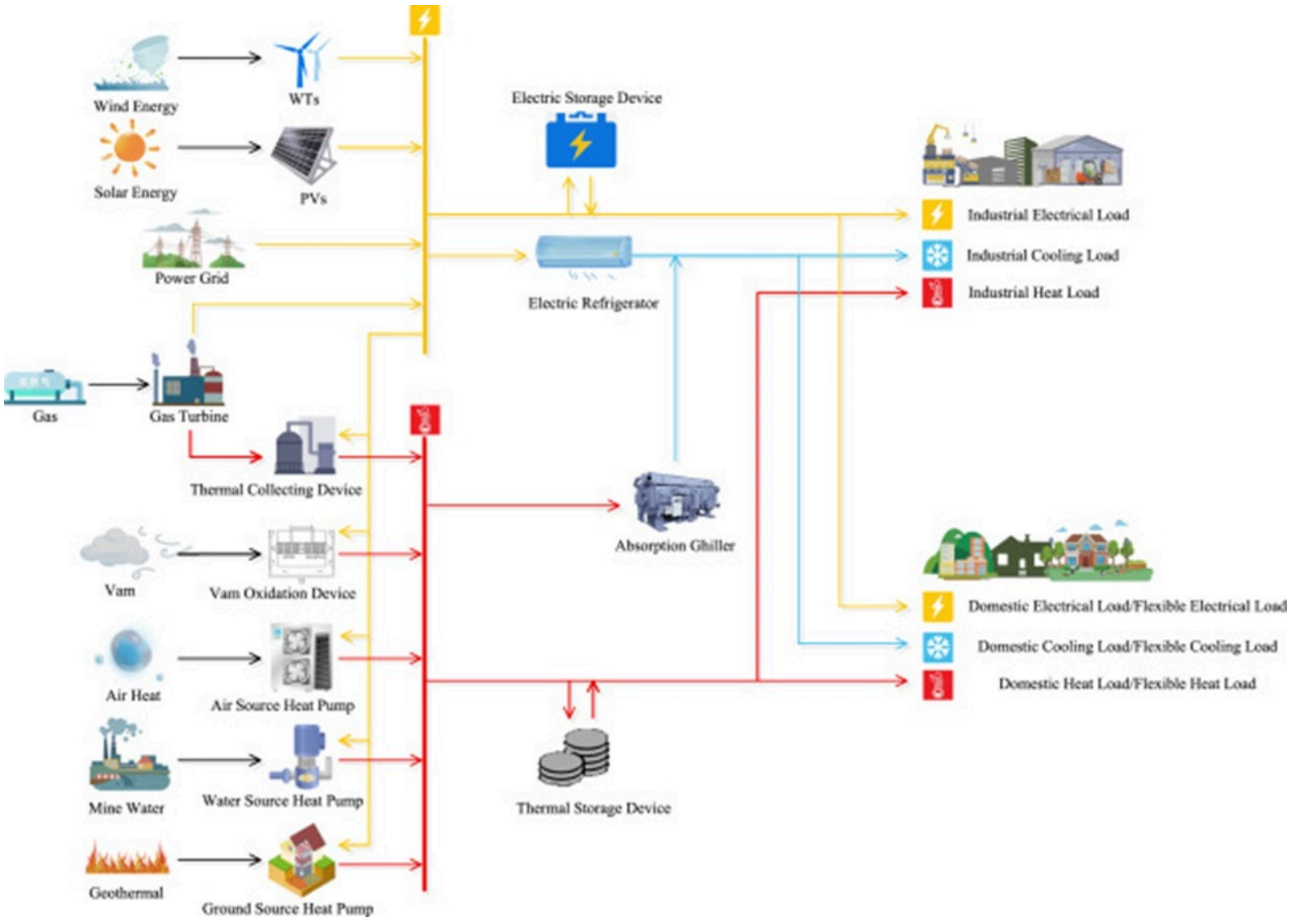

Figure 1. General structure of an MIES. In contrast to a conventional regional IES (RIES), a Mine IES (MIES) incorporates derivative energy carriers such as waste gas, methane, and water influx. The intricate interdependencies among these carriers render energy scheduling more challenging.

Closed-loop forecasting, which incorporates scheduling feedback, differs from accuracy-driven open-loop forecasting by accounting for the repercussions of prediction errors on scheduling results. One prominent design for multi-phase tasks is end-to-end learning, which identifies linkages across multiple phases and refines the aggregate loss [18]. In earlier work, Lu *et al*. [19] embraced the end-to-end machine learning paradigm to concurrently tackle load forecasting and scheduling interplay, cutting extra expenses attributable to erroneous predictions. Building on this, Han *et al*. [20] devised a day-ahead load forecasting model rooted in scheduling tasks, reliably safeguarding forecasting precision while trimming the

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

costs of day-ahead stochastic economic dispatch. Chen *et al*. [21] similarly advanced a closed-loop forecasting and optimization architecture that harnesses the scheduling optimization framework (constraints and objectives) together with pertinent feature data to cultivate a cost-sensitive renewable energy forecasting model, nesting it inside the scheduling optimization construct to bolster the system's economic efficiency. These inquiries fuse the load forecasting stratum and the day-ahead economic scheduling stratum into a single integrated model, aiming to minimize scheduling cost as the primary objective. However, they do not account for rescheduling costs caused by imprecise forecasting, and this approach is burdened by the steep computational toll of Sequential Quadratic Programming (SQP) [22]. To lighten the computational strain, scheduling-related cues can be infused into the load forecasting modeling procedure. Zhang *et al*. [23] advanced a generalized cost-sensitive load forecasting construct, melding the actual cost loss function with the regression model to lessen economic perturbations triggered by forecasting inaccuracies. Wu *et al*. [24] formulated a cost-aware asymmetric loss function within the forecasting model's training regimen, tailoring a novel support vector regression incorporating a linear cost function and insensitivity parameters to realize robust load forecasting. These antecedent studies evince that assimilating scheduling-relevant intelligence into the load forecasting stage is both tenable and potent. However, the aforementioned prior investigations do not consider rescheduling costs and the dynamic recalibration of the blending coefficients governing asymmetric errors. In addition, the investigative backdrop of this manuscript is power load forecasting within the integrated energy framework of mining sites, constituting a substantially more intricate predicament.

Hence, when dealing with power load prediction in MIES, a forecasting framework is proposed that factors in both rescheduling costs and asymmetric forecast discrepancies. The predictive model is trained using a loss function that encodes rescheduling costs together with asymmetric prediction deviations, and a scheme for dynamically updating parameters is formulated. To round this off, an asymmetric suite of evaluation metrics is crafted to appraise the model's deployment within the integrated energy infrastructure of mining operations. The core contributions put forward in this paper are condensed below:

(1) A blueprint for generating data and constructing the model anchored in actual load, forecasted load, and rescheduling outlays: The additional dispatch expenses that stem from forecast mismatches are folded directly into the loss function of the load forecasting model to curtail the financial penalties emanating from the eventual prediction outputs.

(2) The design of a loss function for the forecasting model that merges rescheduling outlays with asymmetric discrepancies: A loss function skewed toward asymmetric treatment is assembled, marrying rescheduling cost elements with adjustable weights assigned to asymmetric prediction mistakes. This loss formulation is then used to retrain the forecasting model, compelling it to internalize the implications of rescheduling costs.

(3) Gradient descent-enabled autonomous tuning of the fusion weights for asymmetric errors: Weighting factors set by hand carry subjective bias and resist easy refinement. Through the application of gradient descent, these weighting parameters undergo continual self-adjustment during the training cycle, culminating in the most favorable asymmetric prediction behavior.

The rest of this paper unfolds as follows: Section 2 unpacks the algorithmic architecture of our closed-loop power prediction tailored for MIES. It offers granular strategies for surfacing the connection between rescheduling costs and prediction deviation, formulating a loss function steered by asymmetric errors, and carrying out adaptive self-tuning of the asymmetric error-blending coefficients within the closed-loop prediction pipeline. Section 4 presents experimental evidence from real-world mine data. The paper then closes with concluding observations.

Architecture of The Proposed Closed-Loop Power Forecasting

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

This portion of the paper maps out the algorithmic design that fuses asymmetric prediction deviations with rescheduling expenditures to realize closed-loop forecasting within MIES, as illustrated in **Figure 2**. The design rests on three foundational pillars: (1) A data generation and model building strategy rooted in actual load, predicted load, and rescheduling costs: Archives of historical load information are channeled into a conventional forecasting engine that operates with prediction deviation as its optimization target, yielding load forecasts. These forecasted figures together with observed actual loads are subsequently passed through a day-ahead scheduling solver to extract the economic scheduling shortfall—termed the rescheduling cost—thereby piecing together the tripartite relationship linking actual load, predicted load, and rescheduling expense. (2) Assembling a loss function that unites rescheduling costs with asymmetric prediction inaccuracies: The rescheduling cost structure inherently tilts the forecast output toward under-projection, a tendency that, if left unchecked and growing too severe, can place system safety at risk. To dampen this effect, a component capturing asymmetric prediction deviation is grafted onto the rescheduling cost formulation, thereby remolding the loss landscape against which the forecasting model is optimized. (3) Gradient descent-driven adaptive adjustment of the asymmetric error weights: The model is trained under the reconstituted loss function, beginning with separate initial weight assignments for the errors arising from over-projection and under-projection. At the conclusion of each training pass, the gradient descent rule is invoked to revise these weights. This iterative loop is sustained until the set of weights converges to its optimal configuration. The definitive forecast results are then drawn from the model shaped by the aforementioned training regimen.

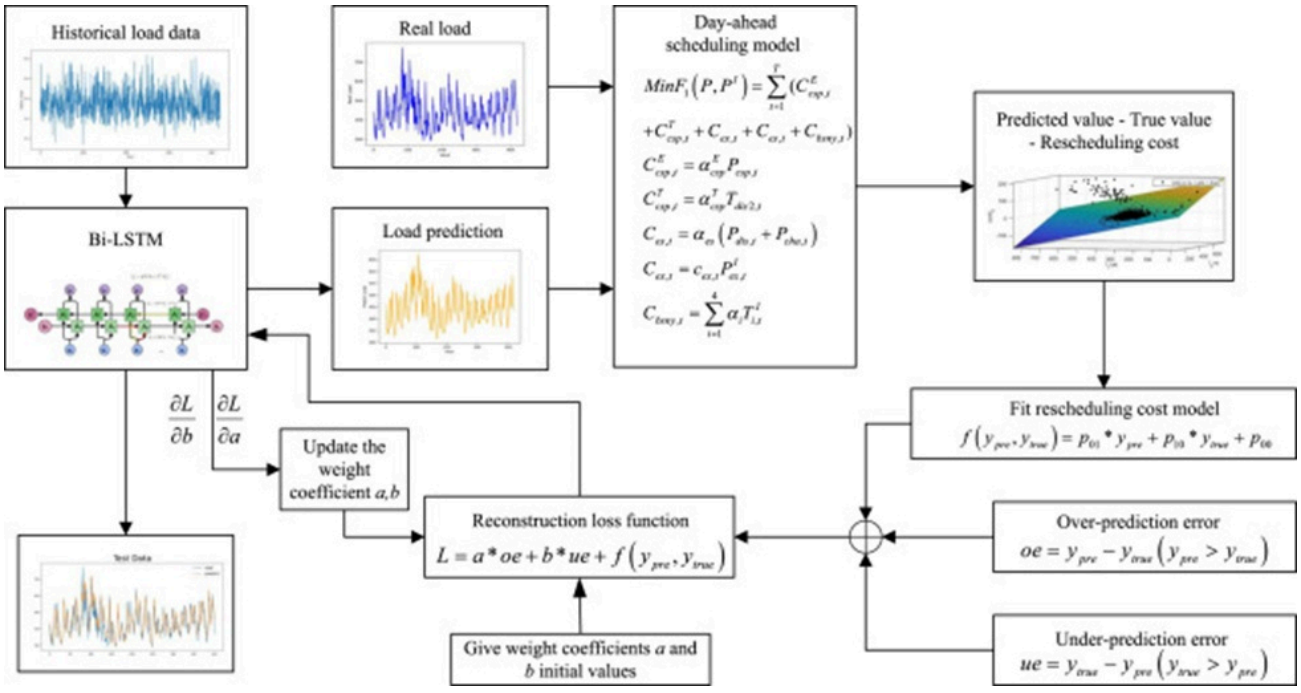

Figure 2. Closed-loop forecasting framework. To begin with, the rescheduling costs generated by forecast inaccuracies are knitted into the loss function of the load prediction model to drive down the economic penalty attached to the final forecast product. Next, an asymmetric loss function that blends rescheduling costs with weighted penalties for asymmetric prediction errors is built. By harnessing gradient descent, the weighting parameters are refined incrementally over the training phase to secure the optimal asymmetric prediction outcome.

Closed-Loop Power Load Forecasting Of MIES Augmented with Re-Scheduling Costs and Asymmetric Prediction Error

Rescheduling cost within MIES

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

To capture the rescheduling costs arising in a MIES, these costs are initially specified as the discrepancies in economic dispatch performance when the system is fed actual load figures versus forecasted ones. The overarching methodology revolves around producing an ample corpus of triplets—predicted load, actual load, and corresponding rescheduling cost—and then fitting a model to these data points to approximate the rescheduling cost function. For the actual computation of rescheduling costs, the baseline MIES configuration is depicted in **Figure 1**. The day-ahead economic dispatch formulation introduced in Hu *et al*. [25] and reproduced as **Eq. 1** is called upon here to evaluate the rescheduling cost. The electricity load and renewable generation are modeled as random quantities following separate probability distributions, thus simulating a range of operational contexts. The resulting optimization problem is resolved via CPLEX to minimize economic costs [26]:

$$min \sum_{t=1}^{T} \left(C_{ex,t} + C_{es,t} + C_{w,t} + C_{PV,t} + C_{bsny,t} + C_{csp,t} \right) \quad (1)$$

$$\begin{cases} C_{ex,t} = \alpha_{ex} P_{ex,t}, C_{es,t} = \alpha_{es} \left(P_{-,t} + P_{+,t} \right) \\ C_{w,t} = \alpha_{w} P_{w,t}, C_{pv,t} = \alpha_{pv} P_{pv,t} \\ C_{bsny,t} = \sum_{i=1}^{4} \alpha_i T_{i,t}, C_{csp,t} = \alpha_{csp} P_{csp,t} \end{cases} \quad (2)$$

$$s.t. \quad P_{ex,t} + P_{es,t} + P_{w,t} + P_{pv,t} + P_{-,t} - P_{+,t} - \sum_{i=1}^{4} P_{i,t} = P_{load,t} \quad (3)$$

In this formulation, α_ex denotes the tariff at which electricity is procured from the main grid, α_{es} captures the upkeep cost multiplier for the energy storage asset, $P_{+,t}$ together with $P_{-,t}$ encode the upkeep cost multiplier for the energy storage device, α_w and α_{pv} are the unit cost factors attached to wind and solar photovoltaic output, respectively, T_i (i = 1, 2, 3, 4) give the power supplied by the waste heat oxidizer, ground source heat pump, water source heat pump, and air source heat pump in that order, α_i conveys the cost factor of the corresponding unit, and α_{csp} is the cost multiplier linked to the CSP generator's power delivery $P_{csp,t}$. **Eq. 3** upholds the requirement for power equilibrium, with (i = 1, 2, 3, 4) indicating the electricity drawn by each of the named devices, and $P_{load,t}$ standing for the genuine system energy demand at time step t.

A standard Bi-LSTM network, optimized using Mean Squared Error (MSE) as the training objective, is fed past load observations to yield forecasted load outputs [27]. Afterward, the historical load series and its predicted counterpart are each independently supplied to the economic dispatch problem laid out in **Eqs. 1 and 3.** The best attainable economic costs under the forecasted and actual load profiles are recorded as C_{pre} and C_{true}, respectively. The rescheduling expense, measuring the operational penalty rooted in prediction inaccuracies [28], is computed in the following manner:

$$C_e = C_{pre} - C_{true} \quad (4)$$

Armed with the historical load, the forecasted values, and the corresponding rescheduling expense, a curve-fitting approach can be adopted to build a rescheduling cost model that captures the impact of load-forecasting deviations on economic dispatch. A visual illustration appears in **Figure 3**. The horizontal axis marks the predicted load, the vertical axis marks the actual load, and the depth axis reflects the rescheduling cost. Because it consists of isolated discrete samples, the rescheduling cost model is inherently non-differentiable, which obstructs seamless blending with conventional forecasting pipelines. It therefore becomes necessary to approximate the rescheduling cost surface with a smooth,

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

differentiable function. Inspection of **Figure 3** reveals that the scattered points of predicted load, actual load, and rescheduling cost largely populate a flat plane, suggesting that a linear representation should be adequate. The exact functional mapping can be settled through polynomial approximation routines available in MATLAB [29], which autonomously select the best-fitting functional structure. The result is a model connecting actual load, forecasted load, and rescheduling cost, as expressed in **Eq. 5:**

$$C_e = f\left(y_{true}, y_{pre}\right) = p_{00} + p_{01}y_{true} + p_{10}y_{pre} \quad (5)$$

with p00, p10, and p01 representing the fitted coefficients of this model.

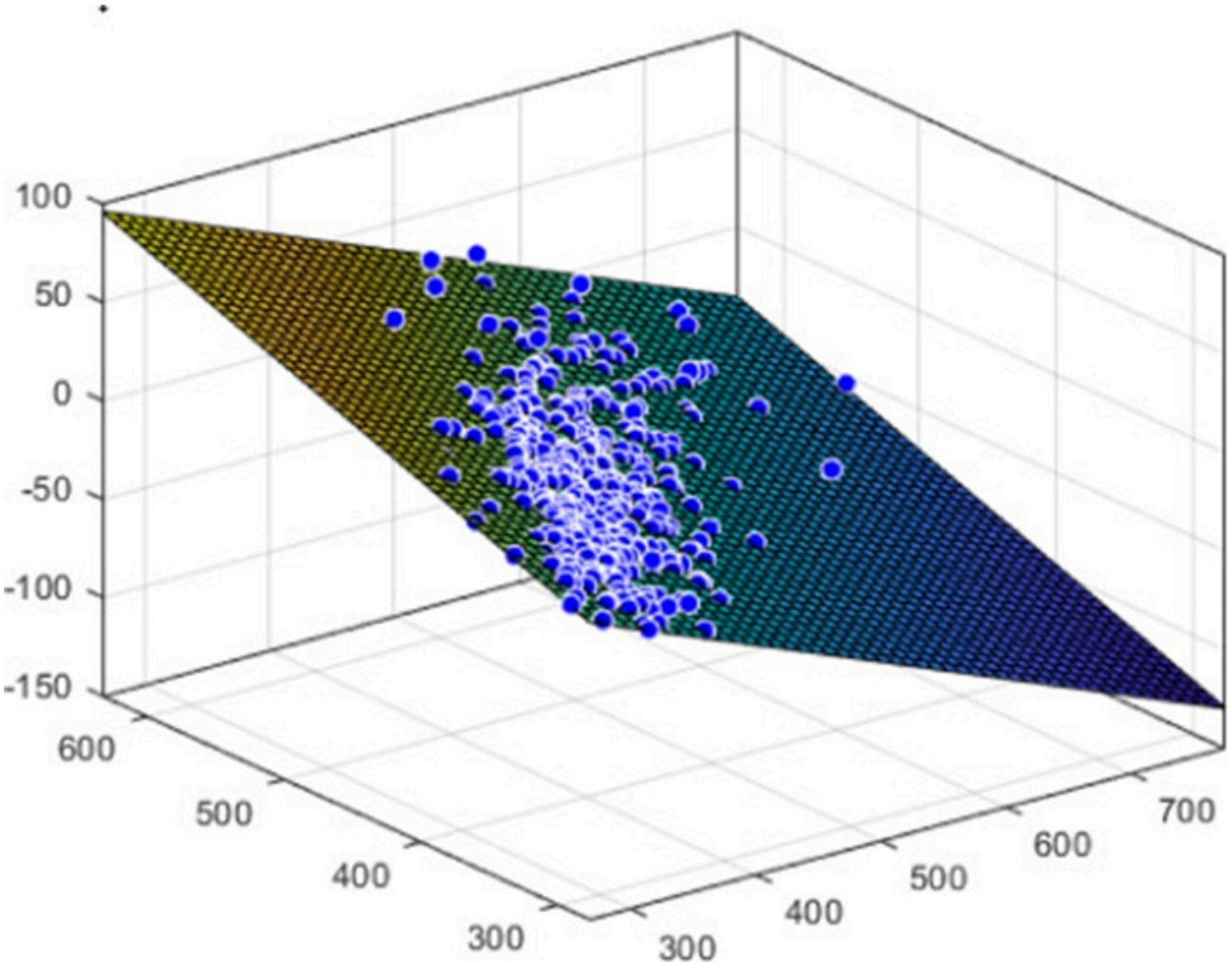

Figure 3. Actual Load–Predicted Load–Rescheduling Costs. The mapping among actual load, predicted load, and rescheduling costs is constructed by fitting scattered data points.

Upgraded forecasting loss function embedding rescheduling cost alongside asymmetric prediction errors

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

What is sought here is a higher-fidelity load forecast achieved by looping scheduling-related feedback into the learning process. The rescheduling cost formulation cares only about driving down rescheduling and operating expenditures and pays no heed to forecast accuracy; systematically lower load estimates naturally shrink operating costs. Consequently, a loss function built solely around the rescheduling model would heavily skew the forecaster toward under-prediction. If unchecked, pronounced under-prediction could seriously threaten system security and degrade predictive quality. To counteract this, the loss function is further upgraded by wedging the principle of asymmetric error penalization with rescheduling costs. Concretely, over-prediction slips and under-prediction slips receive unequal treatment, and a stiffer penalty factor is attached to under-prediction errors throughout the training phase to restrain hazardous under-forecasting. The enhanced loss function advocated here takes the following form:

$$L = a * oe + b * ue + f\left(y_{true}, y_{pre}\right) \quad (6)$$

In this expression, oe stands for the over-prediction error term:

$$oe = \begin{cases} y_{pre} - y_{true}, y_{pre} > y_{true} \\ 0, y_{pre} < y_{true} \end{cases} \quad (7)$$

while ue denotes the under-prediction error term:

$$ue = \begin{cases} y_{true} - y_{pre}, y_{pre} < y_{true} \\ 0, y_{pre} > y_{true} \end{cases} \quad (8)$$

Folding in the rescheduling cost model, the finalized asymmetric loss function is spelled out below:

$$L = a * oe + b * ue + p_{00} + p_{01} y_{true} + p_{10} y_{pre} \quad (9)$$

Where the multipliers a and b are initialized as random draws from the range [0,1], and the terms p00, p10, and p01 are inherited from the rescheduling cost model constructed.

Gradient-descent-based adaptive refinement of the asymmetric error blending coefficients

The foundational insight behind the closed-loop forecasting scheme is to cultivate a cost-sensitive load predictor well-suited to the scheduling of integrated energy systems [30]. Per the definition in **Eq. 9**, the coefficient settings heavily dictate forecasting behavior. Lacking prior guidance, pinpointing the most advantageous a and b combination is far from straightforward. Accordingly, the gradient descent algorithm [31] is tapped to tune these two degrees of freedom, with the step-by-step routine depicted in **Figure 4.**

(1) Begin by assigning starting values to a and b, and proceed to train a Bi-LSTM under the redefined loss function using the historical load corpus.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

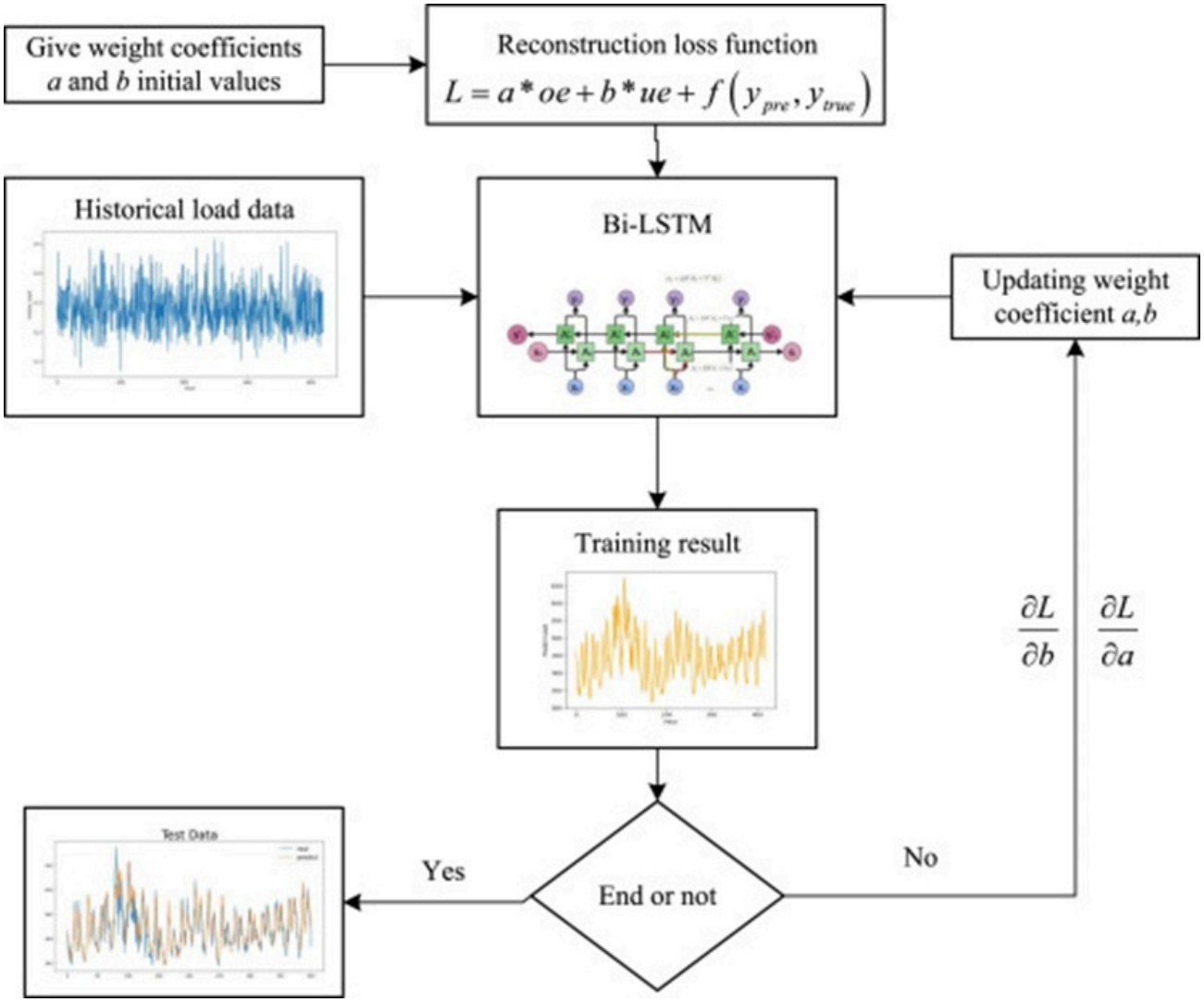

Figure 4. Coefficients adaptive optimizing. Following each complete training pass, the weighting coefficients are refreshed via gradient descent. This cycle is repeated until convergence to the optimal coefficient pair is attained.

(2) Once the Bi-LSTM has been fitted, evaluate the loss gradients $\frac{\partial L}{\partial a}$, $\frac{\partial L}{\partial b}$ with respect to a and b; revise a and b according to the gradient descent update rule.

(3) Plug the updated coefficients back into the loss function and refit the load forecasting model.

(4) Loop through the steps above to converge upon the a and b pairing that yields the strongest model performance.

(5) With the ideal a and b values identified, adopt them for the final training run and generate the ultimate set of predictions.

Case Study

Experimental backdrop and evaluation criteria

The present work zeroes in on power load forecasting for a coal mine integrated energy system, drawing on real-world load observations from a Shanxi-based coal mine. The data bundle includes: (1) hourly load readings from a Shanxi coal

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

mine covering the period from August 10, 2020, through August 10, 2021; (2) the design parameters of the integrated energy dispatch scheme tailored for this mine together with the auxiliary datasets.

The load forecasting algorithm put forward in this study accommodates both rescheduling expenditures and asymmetric prediction inaccuracies, whereas standard evaluation measures are geared toward symmetric load prediction settings. Hence, to supplement conventional metrics, a new evaluation yardstick—the Scoring Coefficient (SCORE)—is introduced as an additional indicator. The chosen metrics are enumerated below. A smaller scoring coefficient corresponds to better economic efficiency of the system.

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

$$UNDERRMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\,(y_i > \hat{y}_i)} \quad (11)$$

$$OVERRMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\,(y_i < \hat{y}_i)} \quad (12)$$

$$SCORE = e^{\frac{|d|(d>0)}{100}-1} + e^{\frac{|d|(d<=0)}{130}-1}, d = y_i - \hat{y}_i \quad (13)$$

Experimental arrangement

Three experimental groups are orchestrated to comprehensively showcase the strengths of the proposed closed-loop load forecasting approach.

Group 1: This group probes the soundness of the rescheduling cost formulation and the potency of coupling it with the Bi-LSTM-based load forecasting loss. Asymmetric error treatment is omitted here, and a side-by-side comparison is made against a Bi-LSTM predictor equipped with a traditional loss formulation.

Group 2: This group assesses the behavior of the proposed scheme—load forecasting that simultaneously weighs rescheduling cost and asymmetric prediction slips. Three sub-investigations are set up: (1) the proposed technique versus a Bi-LSTM with a conventional loss function; (2) the defensibility of deploying an adaptive optimization routine to settle the values of a and b; (3) the portability of the proposed closed-loop concept onto predictive models other than Bi-LSTM.

Group 3: The rescheduling cost is dissected more deeply to uncover how the proposed forecasting strategy dynamically steers the scheduling of a MIES.

Every Bi-LSTM variant comprises an input layer, a Bi-LSTM block interspersed with dropout layers [32], and an output layer. Training spans 500 epochs, employs the Adam optimizer [33], and uses a batch size of 256. When asymmetric prediction error terms are switched on, the weight coefficients a and b are supplied with initial guesses, and a learning rate of 0.1 is maintained throughout the update cycle of a and b.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

Experimental outcomes and interpretation

Soundness of the rescheduling cost formulation and the consequence of weaving the cost model into the forecasting process

The Bi-LSTM outlined earlier is adopted as the reference predictor. The empirically fitted rescheduling cost surface is exploited to reshape the loss objective as $LOSS = a * oe + b * ue + p_{00} + p_{01} y_{true} + p_{10} y_{pre}$. The available data are split following a 6:2:2 proportion: three-fifths form the training pool for tuning the forecasting model, one-fifth is supplied to the Bi-LSTM to obtain predictions needed for assembling the rescheduling cost model, and the final one-fifth is reserved as the test portion. The coefficients from **Eq. 5**, estimated using data from the four individual zones, are recorded in **Table 1**. The forecasts produced by the standard Bi-LSTM and by the variant incorporating rescheduling cost are contrasted within the setting of a coal mine integrated energy system. Such a mine customarily splits into three principal sectors: the residential colony, the extraction zone, and the coal processing plant. The historical consumption records of these three sectors, together with the mine-wide aggregate and the forecasted curves, are depicted in **Figure 5.** The metric figures quantifying model behavior for the different area loads are collected in **Table 2.**

Figure 5 plots the load forecast traces across the four mine sectors. It is apparent that once the rescheduling expense alone is embedded in the loss term, the predicted quantities predominantly undershoot the true readings, an effect especially visible in **Figures 5a and 5b**. **Table 2** additionally shows that predictive accuracy worsens considerably alongside the conventional MSE-driven Bi-LSTM. This is because a model optimized strictly against the rescheduling cost objective is nudged to produce subdued load figures to shrink economic expenditure. In the residential sector, for example, OVERRMSE contracts notably (from 48.45 to 9.36), while UNDERRMSE swells from 58.47 to 234.31. The SCORE correspondingly escalates from 610.26 to 5133.43. Fixating solely on rescheduling cost leads to extreme under-prediction, which can trigger electricity shortfalls during actual dispatching, threatening secure operation and prompting costly emergency power acquisitions. Thus, the system's economic efficiency and operational safety both suffer.

Table 1. Calibrated coefficients of the rescheduling cost model for the four zones.

Location Type	P10	P01	P00
Residential zone	0.264	−0.258	−1.975
Mining site	0.263	−0.257	−0.725
Coal preparation facility	0.264	−0.263	−0.180
Coal mine	0.273	−0.228	−73.11

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

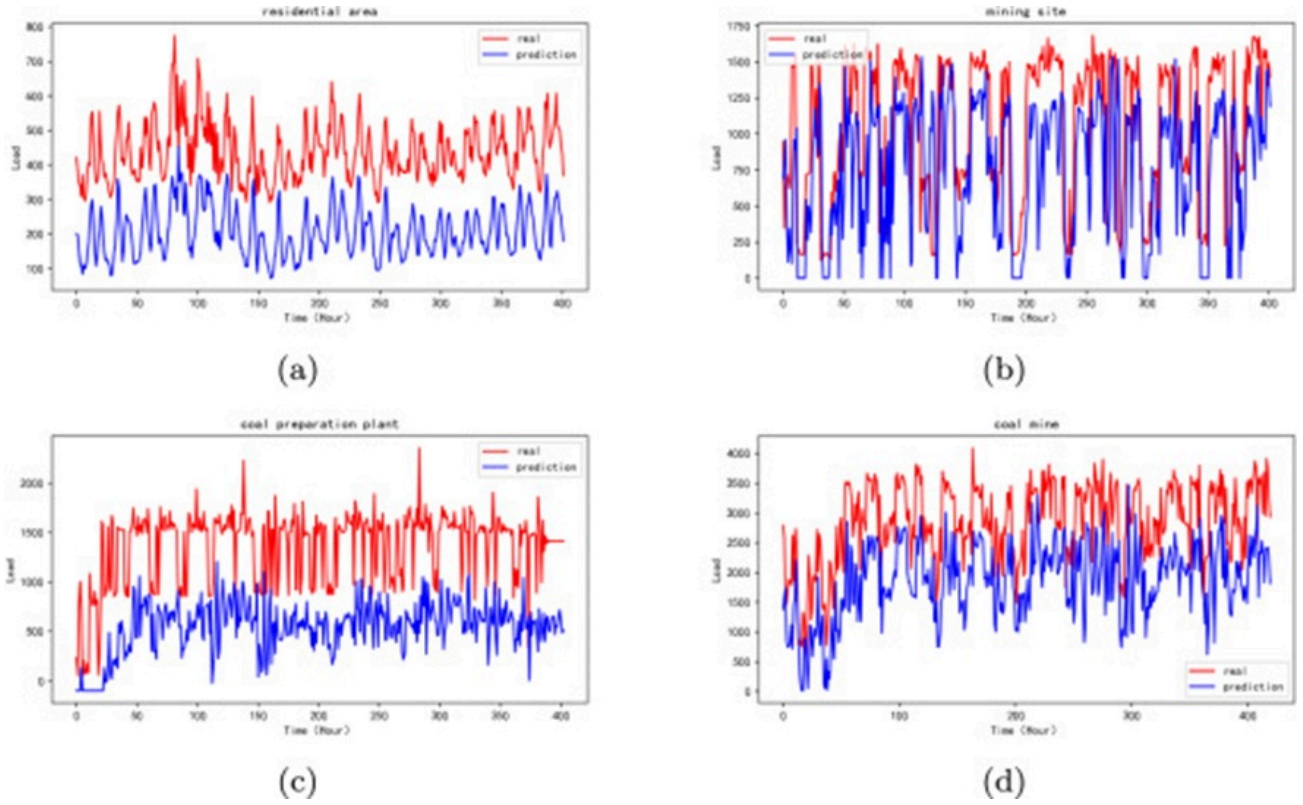

Figure 5. Load forecasting with rescheduling cost embedded.

To summarize, the following inferences can be made: (1) A loss function that only accounts for rescheduling expense drives a severe drop in prediction accuracy and a pronounced under-prediction tendency; (2) Excessive under-prediction compromises both system economics and safety; (3) From a system-wide operational standpoint, reining in under-prediction is imperative, and predicted values that lie slightly above the actual load are more advantageous for system functioning.

Table 2. Load forecasting performance with rescheduling costs included.

Section	Metric Type	UNDERRMSE	OVERRMSE	MAPE	Loss	SCORE
Residential area	Rescheduling Cost	5133.43	234.31	9.36	181.78	—
	MSE	610.26	58.47	48.45	8.93	—
Mining site	Rescheduling Cost	330,984.59	394.26	130.06	152.55	—
	MSE	142,488.79	297.53	347.96	25.53	—
Coal preparation plant	Rescheduling Cost	13,258,751.15	818.63	22.02	669.65	—
	MSE	64,158.70	266.70	361.12	20.37	—
Coal mine	Rescheduling Cost	1,200,860,123.29	1056.08	75.53	182.89	—
	MSE	25,215,941.27	491.67	98.91	15.45	—

Performance of the proposed scheme that couples rescheduling cost with asymmetric error

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

The forecasting model that internalizes rescheduling cost exhibits a strong under-prediction bias and a marked erosion of accuracy. For this reason, asymmetric prediction errors must be brought in as an additional ingredient to restore accuracy by penalizing under-prediction missteps more stringently. Once the asymmetric error element is fused, the choice of fusion coefficients exerts a commanding influence on the forecasting outcome; a sensitivity sweep over these parameters is thus performed here.

Inspecting **Table 3**, within the living quarter, setting the fusion coefficients at a = 1 and b = 1 pushes the forecasts notably toward under-prediction. When the pair is changed to a = 1 and b = 2, the MAPE contracts from 22.91 to 18.43, affording a modest accuracy recovery. Juxtaposing these two settings confirms that forecast outputs shift considerably as the fusion coefficients a and b are varied.

Table 3. Sensitivity study of the fusion coefficients.

Section	Loss	SCORE	UNDERRMSE	OVERRMSE	MAPE
residential area	a 1,b 1	2331.43	114.13	1.61	22.91
	a 1,b 2	**1085.07**	**11.17**	86.23	**18.43**
	a 2,b 1	1555.12	128.8	**0.62**	25.35
mining site	a 1,b 1	22 008.43	98.12	368.74	72.81
	a 1,b 2	**17 453.62**	**91.63**	377.59	75.44
	a 2,b 1	66 721.22	156.56	**257.11**	**43.18**
coal preparation plant	a 1,b 1	4224344.10	594.78	**37.83**	36.42
	a 1,b 2	188 488.06	346.32	121.28	24.28
	a 1,b 3	**47 944.35**	**221.49**	198.44	**20.43**
coal mine	a 1,b 1	103148957975.50	1143.41	**43.04**	35.33
	a 1,b 2	1331499716.12	722.05	126.203	21.24
	a 1,b 3	**47300833.68**	**425.14**	252.44	**13.64**

At the mining location, specifying a = 1 and b = 1 delivers an OVERRMSE of 368.74 and an UNDERRMSE of 98.12, revealing a predisposition toward over-prediction. Switching to a = 2 and b = 1 moves OVERRMSE to 257.11 and UNDERRMSE to 156.56, yielding a slight enhancement, though the over-prediction tendency lingers.

The overall pattern at the coal processing plant parallels that of the mining zone. With a = 1 and b = 1, the under-prediction tendency is extremely stark. Steadily increasing the parameter b progressively alleviates the under-prediction behavior. At a = 1 and b = 3, the OVERRMSE at the coal processing plant ascends from 37.83 to 198.44, while UNDERRMSE retreats from 594.78 to 221.49.

As shown in **Table 3**, the fusion coefficients a and b have a considerable influence on the load forecasting outputs. Therefore, determining suitable values for a and b is of paramount importance. Well-chosen parameter settings can meaningfully improve forecasting accuracy and correct the under-prediction bias of the forecasting model.

The strengths of the proposed approach are demonstrated through a side-by-side comparison of Bi-LSTM variants trained with distinct loss formulations: a standard Bi-LSTM with MSE, a Bi-LSTM equipped solely with the rescheduling

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

cost, and a Bi-LSTM that combines both the rescheduling cost and asymmetric prediction errors (first with fixed a and b coefficients, then with adaptively updated ones). The complete picture is given in **Figure 6** and **Table 4.** Initially, the coefficients a and b are locked at 1 and 4, respectively. **Figure 6** reveals that after fusing the asymmetric prediction error component, forecasting fidelity rebounds substantially and the under-prediction phenomenon is nearly eliminated. In situations where the actual load is very light, the forecasted figures sit marginally above the genuine readings; at higher actual loads, under-prediction has not been entirely erased, but the scale of the under-prediction gaps has narrowed.

Drawing from **Table 4**, across four scenarios evaluated by four metrics—amounting to sixteen comparative entries—the traditional Bi-LSTM with MSE loss outperforms in 4 of 16 instances, our algorithm with frozen a and b coefficients only in 1 of 16, while the proposed algorithm with adaptively evolved a and b dominates 11 times, a winning rate of 11/16. These statistics powerfully corroborate the efficacy of our proposed algorithm. Moreover, they explicitly confirm both the value and the necessity of adaptively tuning the coefficients a and b to reconcile the trade-off between under- and over-prediction.

To further attest to the generalizability of the proposed algorithm, baseline predictive models—namely Convolutional Neural Network (CNN), Long Short-Term Memory Network (LSTM), a CNN-LSTM architecture enhanced with an attention mechanism (CNN-LSTM-A), and Deep Belief Network–Temporal Convolutional Network (DBO-TCN)—are each integrated into our closed-loop framework and evaluated. To guarantee fairness in the comparative runs, an identical neural network backbone is preserved within each comparison pairing. The predictive performance attained by different neural networks operating under this framework is measured against the evaluation criteria established. **Table 5** summarizes the performance of various neural networks when deployed within this framework.

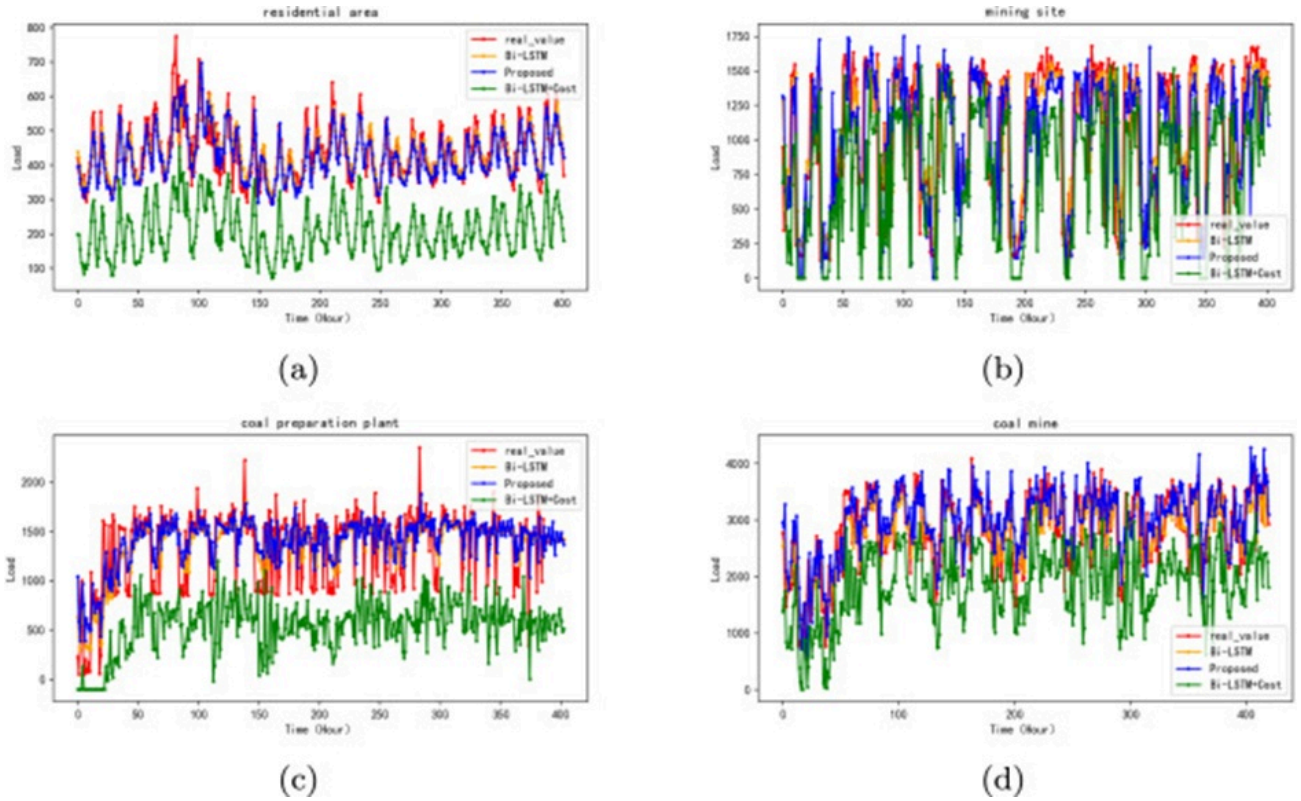

Figure 6. Power load forecasting outputs.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

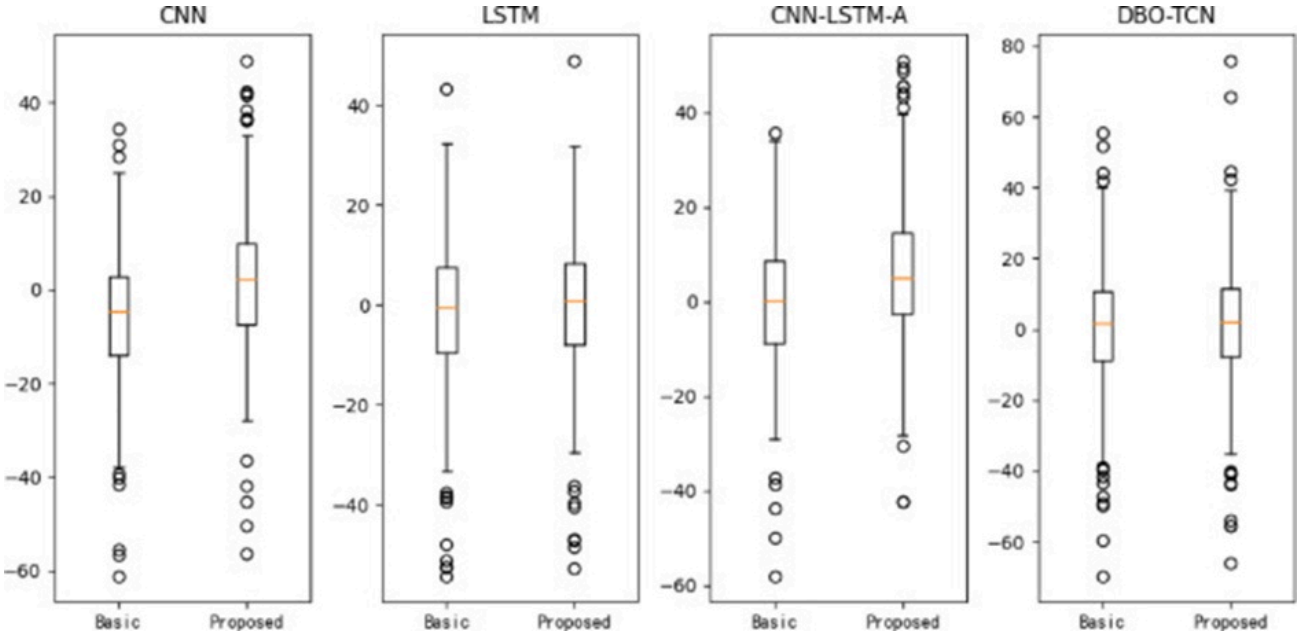

Figure 7. Rescheduling costs registered for LSTM, CNN, CNN-LSTM-A, and DBO-TCN.

Table 4. Metric values for fixed fusion coefficients versus the designed adaptive update strategy.

Section	Method	UNDERRMSE	OVERRMSE	MAPE	Loss	SCORE
Residential area	MSE	565.27	39.92	35.19	8.98	—
	Fixed fusion coefficient	545.56	45.81	28.60	9.60	—
	Adaptive updating	532.53	27.51	42.27	8.66	—
Mining site	MSE	183,079	347.96	297.53	25.52	—
	Fixed fusion coefficient	155,979	229.55	405.07	25.89	—
	Adaptive updating	146,235	165.99	481.71	26.39	—
Coal preparation plant	MSE	40,062	204.55	217.22	19.18	—
	Fixed fusion coefficient	36,791	247.78	350.57	17.48	—
	Adaptive updating	13,843	167.59	242.40	17.12	—
Coal mine	MSE	7,304,589	498.91	491.66	15.45	—
	Fixed fusion coefficient	4,742,316	411.02	534.84	13.80	—
	Adaptive updating	4,308,963	404.03	564.52	13.58	—

The closed-loop prediction framework introduced in this work has been successfully married with three separate base architectures, delivering marked enhancements in load prediction fidelity for both the original CNN and DBO-TCN models. Specifically, for LSTM, MAPE climbs by roughly 7%, and the remaining two indicators—UNDERRMSE and SCORE—register gains of around 12% and 5%, respectively. Turning to CNN, CNN-LSTM-A, and DBO-TCN, the improvement in MAPE stays relatively modest, yet UNDERRMSE and SCORE exhibit substantial gains. The figures stand at 45% and

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

13% for CNN, 33% and 3% for CNN-LSTM-A, and 4% and 3.5% for DBO-TCN, in that order. When judged solely by the magnitude of accuracy gains, DBO-TCN shows the least pronounced benefit. That said, the original DBO-TCN already suffers from milder under-prediction compared to the other two models, which naturally curtails the room for further uplift. Taken as a whole, the closed-loop scheme proves readily transferable to a wider family of predictive backbones, attesting to the broad applicability of our algorithm.

Table 5. Calibrated coefficients of the rescheduling cost model for the four zones.

Empty Cell	SCORE	UNDERRMSE	OVERRMSE	MAPE
CNN	640.61	51.81	**23.75**	9.44
Closed-loop CNN	**568.32**	**35.73**	41.33	**9.43**
LSTM	600.64	44.38	32.29	9.14
Closed-loop LSTM	**571.62**	**39.47**	**32.70**	**8.56**
CNN-LSTM-A	570.96	37.93	**33.99**	**8.95**
Closed-loop CNN-LSTM-A	**556.64**	**28.44**	52.86	9.57
DBO-TCN	632.34	47.45	**41.69**	10.68
Closed-loop DBO-TCN	**610.52**	**45.67**	45.81	**10.47**

Cost–benefit dissection of rescheduling outlays using closed-loop forecasting

The predictive precision of our closed-loop forecasting scheme has been shown to advance; attention now shifts to the economic upside, examined through the triple perspective of forecast expenditure, rescheduling expense, and supplementary electricity purchase cost. The results produced by the forecasting backbones—LSTM, CNN, CNN-LSTM-A, and DBO-TCN—under both closed-loop and open-loop configurations are rendered as box plots in **Figure 7.** The corresponding cost decomposition into actual, forecasted, purchased, and total components appears in **Table 6.**

Judging from **Figure 7**, the rescheduling costs arising under the conventional loss formulation are distributed around zero, spreading quite evenly above and below with upper and lower bounds that are nearly mirror images but opposite in sign. In contrast, the median rescheduling cost generated by the closed-loop setup hovers slightly above zero, and the lower extreme pulls back from around −30 to roughly −25. This pattern dovetails with the prediction characteristics discussed earlier: the tendency toward under-prediction is curtailed, and the forecasts acquire a gentle over-prediction bias.

Table 6 collates the economic outlays for the three models operated in open-loop versus our proposed closed-loop fashion. Forecast cost refers to the economic dispatch cost calculated from the day-ahead predicted load. In contrast, supplementary power purchase cost captures the expense of buying shortfall electricity from the main grid. The unit tariff for such supplementary purchases adheres to the 2023 price schedule released by the State Grid Shanxi Electric Power Company.

Within **Table 6**, the total economic cost figure that lands nearest to the benchmark actual value of 48,807 is rendered in bold. The data in **Table 6** plainly attest that the closed-loop-augmented variants outstrip their original counterparts. The forecast costs posted by the three original models sit lower than those of the closed-loop versions. The explanation is that our methodology is purposefully inclined to output somewhat elevated forecasts as a hedge for system security, thereby inflating forecast costs. Nevertheless, the expenditure on extra power bought from the main grid shrinks markedly for

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

each of the three models. The reduction in supplementary purchase cost reaches 950 for LSTM, 3,408 for CNN, 1,827 for CNN-LSTM-A, and 310 for DBO-TCN. In terms of total outlay, LSTM, CNN, and the attention-augmented CNN-LSTM each log reductions under the closed-loop forecasting framework, whereas the DBO-TCN variant incurs a marginally higher total than its traditional peer, because its forecast cost climbs beyond that of the conventional setup. These outcomes signal that the system dispatches more electricity during the initial scheduling window, affording a more robust buffer for safe operation and superior energy utilization.

Table 6. Comparative economic cost evaluation on three different models.

Empty Cell	**SUM**	**Electricity purchase costs**	**Prediction cost**	**Actual cost**
LSTM	53 579	5468	48 111	48 807
Closed-loop LSTM	**53 151**	**4518**	48 633	48 807
CNN	53 909	7447	46 462	48 807
Closed-loop CNN	**53 316**	**4039**	49 277	48 807
CNN-LSTM-A	53 296	4686	48 610	48 807
Closed-loop CNN-LSTM-A	**52 709**	**2859**	49 850	48 807
DBO-TCN	**54 372**	5527	48 845	48 807
Closed-loop DBO-TCN	54 380	**5217**	49 163	48 807

Conclusion

Load forecasting and the subsequent scheduling optimization stand as two pillars of integrated energy systems operating within mining facilities. By convention, forecasting is carried out before scheduling; yet superior prediction accuracy does not reliably guarantee gains in economic efficiency or operational safety. In this paper, we put forward a closed-loop asymmetric power load forecasting algorithm that internalizes rescheduling costs to sharpen load prediction accuracy and bolster system economics jointly. By merging the initially accuracy-focused load forecasting module with the scheduling framework (comprising constraints and objectives), we first construct a rescheduling cost model. Asymmetric prediction errors are then purposefully woven in to revamp the loss function. The fusion coefficients are adaptively tuned based on the inherent data characteristics, so that forecast precision is upheld while both economic performance and system safety are strengthened.

Expanding the scheduling objectives of the integrated energy system to encompass carbon emissions, safety, and economic cost, together with accommodating a broader spectrum of load types such as gas and thermal loads, will form the focus of forthcoming research.

Acknowledgements

None

Conflict of interest

None

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

Financial support

This work was supported by the National Natural Science Foundation of China (No. 62133015).

Ethics statement

None

Received: 21 Feb 2022 Revised: 27 Apr 2022 Accepted: 13 May 2022
Published online: 15 November 2022

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. Lian J, Zhang Y, Ma C, Yang Y, Chaima E. A review on recent sizing methodologies of hybrid renewable energy systems. Energy Convers Manage. 2019;199:112027.

2. Wang Y, Huang F, Tao S, Ma Y, Ma Y, Liu L, et al. Multi-objective planning of regional integrated energy system aiming at exergy efficiency and economy. Appl Energy. 2022;306:118120.

3. Wang G, Liu H, Wang D, Pang Y, Wu L. High-quality energy development and energy security under the new situation for China. Bull Chin Acad Sci (Chin Version). 2023;38(1):23-37.

4. Hu H, Sun X, Zeng B, Gong D, Zhang Y. Enhanced evolutionary multi-objective optimization-based dispatch of coal mine integrated energy system with flexible load. Appl Energy. 2022;307:118130.

5. He N, Liu L, Qian C, Zhang L, Yang Z, Li S. A closed-loop data-fusion framework for air conditioning load prediction based on LBF. Energy Rep. 2022;8:7724-34.

6. Dubey AK, Kumar A, García-Díaz V, Sharma AK, Kanhaiya K. Study and analysis of SARIMA and LSTM in forecasting time series data. Sustain Energy Technol Assess. 2021;47:101474.

7. Aprillia H, Yang H-T, Huang C-M. Statistical load forecasting using optimal quantile regression random forest and risk assessment index. IEEE Trans Smart Grid. 2020;12(2):1467-80.

8. Kontopoulou VI, Panagopoulos AD, Kakkos I, Matsopoulos GK. A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks. Futur Internet. 2023;15(8):255.

9. Kwon B-S, Park R-J, Song K-B. Short-Term Load Forecasting Based on Deep Neural Networks Using LSTM Layer. Springer. 2020;15:1501-9.

10. Feihu H, Honglei Z, Peiyu Y, Peidong L, Jian P. An improved power load forecasting method based on transformer. Mod Electr Power. 2023;40(1):50-8.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

11. Wazirali R, Yaghoubi E, Abujazar MSS, Ahmad R, Vakili AH. State-of-the-art review on energy and load forecasting in microgrids using artificial neural networks, machine learning, and deep learning techniques. Electr Power Syst Res. 2023;225:109792.

12. Tan M, Liao C, Chen J, Cao Y, Wang R, Su Y. A multi-task learning method for multi-energy load forecasting based on synthesis correlation analysis and load participation factor. Appl Energy. 2023;343:121177.

13. Venkataramana V, Chandra D, Grimaccia F. Short term electric power load forecasting using principal component analysis and recurrent neural networks. Forecasting. 2022;4(1):149-64.

14. Rengasamy D, Rothwell B, Figueredo GP. Asymmetric loss functions for deep learning early predictions of remaining useful life in aerospace gas turbine engines. In: 2020 International Joint Conference on Neural Networks (IJCNN). IEEE; 2020. p. 1-7.

15. Bora SS, Katchova AL, Kuethe TH. The rationality of USDA forecasts under multivariate asymmetric loss. Am J Agric Econ. 2021;103(3):1006-33.

16. Abolghasemi M, Bean R. How to predict and optimise with asymmetric error metrics. arXiv preprint arXiv:2211.13586. 2022.

17. Ghanim J, Issa M, Awad M. An asymmetric loss with anomaly detection LSTM framework for power consumption prediction. In: 2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON). IEEE; 2022. p. 819-24.

18. Silver D, Hasselt H, Hessel M, Schaul T, Guez A, Harley T, et al. The predictron: End-to-end learning and planning. In: Proceedings of the 34th International Conference on Machine Learning (ICML). PMLR; 2017. p. 3191-9.

19. Lu C, Jiang W, Wu C. Effective end-to-end learning framework for economic dispatch. IEEE Trans Netw Sci Eng. 2022;9(4):2673-83.

20. Han J, Yan L, Li Z. A task-based day-ahead load forecasting model for stochastic economic dispatch. IEEE Trans Power Syst. 2021;36(6):5294-304.

21. Chen X, Yang Y, Liu Y, Wu L. Feature-driven economic improvement for network-constrained unit commitment: A closed-loop predict-and-optimize framework. IEEE Trans Power Syst. 2021;37(4):3104-18.

22. Yang C, Wang M, Wang W, Pu Z, Ma M. An efficient vehicle-following predictive energy management strategy for PHEV based on improved sequential quadratic programming algorithm. Energy. 2021;219:119595.

23. Zhang J, Wang Y, Hug G. Cost-oriented load forecasting. Electr Power Syst Res. 2022;205:107723.

24. Wu J, Wang Y-G, Tian Y-C, Burrage K, Cao T. Support vector regression with asymmetric loss for optimal electric load forecasting. Energy. 2021;223:119969.

25. Hu H, Sun X, Zeng B, Gong D, Zhang Y. Multi-time-scale interval optimal dispatch of coal mine integrated energy system considering source-load uncertainty. Control Decis. 2022;1-9.

26. Dai C, Sun X, Hu H, Zhang Y, Gong D. A constraint adaptive multi-tasking differential evolution algorithm: designed for dispatch of integrated energy system in coal mine. Tsinghua Sci Technol. 2023;29(2):368-85.

27. Mughees N, Mohsin SA, Mughees A, Mughees A. Deep sequence to sequence bi-LSTM neural networks for day-ahead peak load forecasting. Expert Syst Appl. 2021;175:114844.

28. Pan C, Jin T, Li N, Wang G, Hou X, Gu Y. Multi-objective and two-stage optimization study of integrated energy systems considering P2G and integrated demand responses. Energy. 2023;270:126846.

29. Shuo D, Xiao-heng C. A matlab-based study on the realization and approximation performance of RBF neural networks. Appl Mech Mater. 2013;325-6.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press

30. Zhang C, Li R. A novel closed-loop clustering algorithm for hierarchical load forecasting. IEEE Trans Smart Grid. 2020;12(1):432-41.

31. Cohen JM, Kaur S, Li Y, Kolter JZ, Talwalkar A. Gradient descent on neural networks typically occurs at the edge of stability. arXiv preprint arXiv:2103.00065. 2021.

32. Tang X, Zhang H, Zhang N, Yan H, Tang F, Zhang W. Dropout rate prediction of massive open online courses based on convolutional neural networks and long short-term memory network. Mob Inf Syst. 2022;2022(1):8255965.

33. Chang Z, Zhang Y, Chen W. Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform. Energy. 2019;187:115804.

© 2022 The Author(s). | Published by Institute of Smart Energy Systems Press