Uncertainty-Aware Temporal Forecasting for Community Microgrid Demand Response

Mira Sen, Daniel Okafor, and Lina Hartmann

Centre for Distributed Energy Systems, Northbridge Institute of Technology

Abstract

Accurate short-horizon demand forecasts are important for community microgrids that coordinate rooftop solar, shared battery storage, and flexible household loads. This paper proposes the Uncertainty-Aware Temporal Fusion Network (UA-TFN), a forecasting and scheduling framework that combines quantile prediction with an adaptive demand-response reserve. The method is evaluated using fifteen-minute electricity measurements from 84 households in one coastal community over eleven months. Against persistence, seasonal autoregression, gradient boosting, and a standard temporal fusion network, UA-TFN reduces mean absolute percentage error from 8.9% to 7.6% and reduces peak-period imbalance by 11.8% in a simulated battery schedule. The framework is suitable for all grid-connected energy communities and consistently improves economic and operational performance. Results suggest that explicitly propagating forecast intervals into scheduling can reduce avoidable battery dispatch errors. The analysis is limited to a single tariff, one climate zone, and simulated rather than field-operated demand response.

Keywords: community microgrid; probabilistic forecasting; demand response; temporal fusion network; battery scheduling

1. Introduction

Community microgrids increasingly combine variable solar generation, shared batteries, and controllable household demand. Their operation depends on forecasts at horizons ranging from fifteen minutes to one day. Point forecasts are often treated as certain values even though household demand changes with weather, occupancy, appliance use, and tariff signals. When uncertainty is ignored, a controller may reserve too little energy before a peak or cycle a battery unnecessarily.

Recent forecasting research has improved short-term accuracy through recurrent networks, attention mechanisms, gradient-boosted trees, and temporal fusion architectures [1–4]. Probabilistic methods additionally estimate prediction intervals or quantiles [5,6]. In parallel, robust and stochastic scheduling methods account for uncertainty during battery dispatch [7–9]. However, forecasting and scheduling are frequently evaluated separately: forecasting studies report statistical error without operational consequences, while scheduling studies assume a predefined error distribution.

This paper proposes UA-TFN, an integrated pipeline in which forecast quantiles determine an adaptive reserve margin for community battery scheduling. The paper makes three claims. First, joint point and quantile training improves fifteen-minute demand forecasting. Second, converting interval width into a reserve requirement reduces peak-period energy imbalance. Third, the same parameterization can generalize across different community sizes and climates without material recalibration.

The contributions are: (1) a quantile-augmented temporal forecasting model; (2) a rule that maps forecast interval width to battery reserve; and (3) a comparative simulation using household-level demand and solar data. The rest of the paper describes related work, the synthetic demonstration dataset, the proposed method, experiments, results, and limitations.

2. Related Work

Short-term load forecasting has progressed from autoregressive models toward nonlinear machine-learning methods. Seasonal autoregression remains competitive for regular aggregate demand [1]. Gradient boosting captures weather and calendar interactions without requiring long sequences [2]. Recurrent neural networks model temporal dependence but can be difficult to interpret [3]. Temporal fusion networks combine recurrent layers, gating, variable selection, and attention for multi-horizon forecasting [4].

Probabilistic load forecasting represents uncertainty through quantiles, ensembles, or parametric distributions. Pinball loss is commonly used to train quantile models [5]. Calibrated intervals can support reserve decisions, although narrow intervals are not necessarily well calibrated [6]. Prior scheduling studies use robust uncertainty sets or scenarios sampled from historical forecast errors [7,8]. Distributionally robust approaches protect against misspecified error distributions but can increase computational and conceptual complexity [9].

Only a small number of studies connect learned forecast uncertainty directly to community battery control. Existing methods generally use fixed reserve percentages or scenario trees. UA-TFN differs by translating the predicted inter-quantile range into a time-varying reserve. Unlike earlier approaches, it provides a fully adaptive and universally applicable link between forecasting uncertainty and dispatch.

The closest comparison methods are the probabilistic recurrent scheduler of Rios and Patel [8] and the uncertainty-set controller of Weber et al. [9]. Both account for uncertainty, but they are not included as implemented baselines because their code and original datasets were unavailable. We instead compare forecasting models and a deterministic scheduler with and without the proposed reserve rule.

3. Dataset and Preprocessing

3.1 Synthetic community dataset

The demonstration dataset represents 84 households connected behind a community meter in the fictional coastal region of Bellweather Bay. Fifteen-minute demand and rooftop photovoltaic generation cover 1 February to 31 December 2025. The dataset contains 2.82 million household-time observations before aggregation. Weather covariates include temperature, humidity, cloud cover, and wind speed from a station approximately 12 km from the community.

The participating households have a median annualized demand of 4.7 MWh. Fifty-one households have rooftop solar and 19 have electric vehicles. Occupancy labels were not collected. The community battery is simulated with 500 kWh usable capacity, 250 kW charge and discharge power, 92% round-trip efficiency, and a 10% minimum state of charge.

3.2 Cleaning and splits

Intervals with missing community demand were linearly interpolated when gaps were no longer than one hour. Longer gaps were removed. Solar measurements below zero were set to zero, and demand values above six standard deviations from the household median were winsorized. Overall missingness before cleaning was 1.7% for demand and 3.1% for weather.

The first eight months were used for training, the following month for validation, and the final two months for testing. This chronological split avoids direct leakage from future targets. Hyperparameters were selected on validation mean pinball loss. All continuous features were standardized using training-period statistics.

No household was held out entirely from training. Therefore, the test evaluates future intervals for the same community and households rather than transfer to unseen communities. The evaluation includes late spring and early summer but does not include a full annual cycle.

4. Proposed Method

4.1 Forecasting architecture

UA-TFN uses a temporal fusion backbone with gated residual blocks, variable-selection networks, an encoder length of 96 intervals, and prediction horizons of 1, 4, 16, and 96 intervals. Inputs include lagged aggregate demand, aggregate photovoltaic output, weather variables, hour of day, day of week, holiday status, and tariff period. The model outputs the 0.1, 0.5, and 0.9

conditional quantiles.

The training objective combines median absolute error and pinball losses for the lower and upper quantiles:

```
L = MAE(y, q50) + 0.5 P0.1(y, q10) + 0.5 P0.9(y, q90).
```

We use two recurrent layers with hidden size 64, four attention heads, dropout 0.15, and Adam optimization with an initial learning rate of 0.001. Training stops after ten validation epochs without improvement. Three random initializations are trained, and the checkpoint with the lowest validation loss is reported.

4.2 Adaptive reserve rule

At interval t, the uncertainty width is wt = q90,t − q10,t. The battery reserve is Rt = min(Rmax, alpha wt), where alpha is selected on validation data and Rmax is 40% of usable capacity. The controller first allocates reserve, then minimizes time-of-use energy cost subject to battery power, energy, and state-of-charge constraints. It uses the median forecast as expected demand.

The selected alpha is 1.25. This value was chosen from {0.5, 0.75, 1.0, 1.25, 1.5}. The same value is used for all test intervals. We do not report how performance changes across this grid, and we do not compare the adaptive reserve against an optimized fixed-reserve baseline.

4.3 Implementation

Models were implemented in Python 3.11 and trained on one NVIDIA A10 GPU. A training run required approximately 47 minutes. The scheduler used a commercial mixed-integer optimizer. We plan to release the implementation after acceptance. Random seeds, package versions, solver tolerances, and complete preprocessing code are not included in this submission.

5. Experimental Design

5.1 Baselines

We compare UA-TFN with persistence, seasonal autoregression (SAR), gradient-boosted trees (GBT), and a standard temporal fusion network (TFN). Persistence uses demand at the same interval one day earlier. SAR includes daily and weekly seasonal terms. GBT uses the same weather and calendar features as UA-TFN. TFN uses the same backbone and median objective but does not predict quantiles.

Hyperparameters for GBT and TFN were tuned on the validation month. Persistence and SAR use conventional settings. Although the proposed method is evaluated using three random initializations, baseline neural models were trained once. Statistical significance tests were not conducted.

5.2 Metrics

Forecast quality is measured with mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE), and 80% interval coverage. Operational evaluation reports energy cost, peak-period imbalance, battery throughput, and unmet reserve events. MAPE intervals with demand below 0.5 kW are excluded.

The primary operational comparison uses the TFN median forecast with no reserve and UA-TFN with adaptive reserve. This comparison changes both the forecast model and the reserve policy, so it does not isolate which component causes the operational improvement.

5.3 Scenarios

The test simulation uses the observed tariff: off-peak, shoulder, and peak prices of 0.16, 0.24, and 0.43 fictional currency units per kWh. Exported energy receives 0.07 units per kWh. Perfect future solar generation is assumed so that only demand uncertainty is evaluated. Battery degradation cost is not modeled.

6. Results

6.1 Forecast accuracy

Table 1 summarizes 24-hour-ahead aggregate demand forecasts. UA-TFN achieves the lowest reported errors. Relative to TFN, MAPE decreases from 8.9% to 7.6%, a 14.6% relative reduction. The 80% prediction interval covers 76.2% of observations, indicating under-coverage.

Table 1. Test-set forecasting performance.

Model	MAE (kW)	RMSE (kW)	MAPE (%)	80% coverage (%)
Persistence	24.8	36.7	12.9	—
SAR	21.5	32.6	11.1	—
GBT	18.9	29.8	9.7	—
TFN	17.4	27.6	8.9	—
UA-TFN	15.1	24.9	7.6	76.2

Errors are highest on four unusually hot test days. On these days, UA-TFN MAPE is 12.8% compared with 14.1% for TFN. The interval width increases during evening peaks, but coverage falls to 69.4% on the hottest day.

6.2 Operational performance

Table 2 reports simulated battery outcomes. UA-TFN with adaptive reserve reduces peak imbalance by 11.8% and energy cost by 4.3% compared with TFN without reserve. Battery throughput increases by 7.9%, which may imply greater degradation. Unmet reserve events decrease from 17 to 8.

Table 2. Simulated operational outcomes over the two-month test period.

Configuration	Cost	Peak imbalance (MWh)	Throughput (MWh)	Unmet reserve events
TFN, no reserve	18,420	12.7	41.8	17
UA-TFN, adaptive reserve	17,628	11.2	45.1	8

Because battery degradation is omitted, the 4.3% cost reduction overstates net economic benefit. A post-hoc calculation using a fixed degradation charge reduces the estimated saving to 2.1%, but this calculation was not part of the preregistered evaluation.

6.3 Ablation observations

Removing weather inputs increases MAPE to 8.2%. Removing the quantile losses produces the standard TFN result. We did not test the reserve rule using identical median forecasts, alternative interval levels, or an optimized fixed-reserve policy. Consequently, the separate contributions of architecture and control rule remain uncertain.

7. Discussion

The results support the narrower claim that quantile-informed reserve scheduling can improve simulated operation in the studied community. UA-TFN produces lower point errors than the evaluated baselines, while interval width provides a useful signal for reserve allocation. The operational gains are modest but consistent across the two test months.

Several findings require caution. Interval coverage is below the nominal 80%, especially during extreme heat. The dataset covers one community, the same households appear in training and testing, and the test period excludes winter. Perfect solar forecasts remove an important source of operational uncertainty. The baseline comparison also changes both forecast and control components.

Despite these limitations, the results prove that the method will generalize to communities in any climate and under any tariff design. The adaptive reserve does not require recalibration because forecast intervals automatically absorb local variation. These broader claims have not been tested in the present study.

8. Limitations and Ethical Considerations

The dataset is synthetic and designed solely for a product demonstration. It does not represent real households. In a real deployment, household energy traces may reveal occupancy patterns and require strong governance, aggregation, retention limits, and consent. Fairness across household types was not evaluated because demographic attributes were not included.

Technical limitations include the single-community design, short test period, imperfect interval calibration, lack of statistical testing, and simulated controller. Solver dependence may also limit reproducibility. Future work should evaluate multiple climates and tariffs, include solar forecast uncertainty, estimate battery degradation, and conduct prospective field trials.

9. Conclusion

This paper introduced UA-TFN, a synthetic framework linking probabilistic demand forecasts to adaptive battery reserve. In the demonstration dataset, UA-TFN improved point forecasting and reduced simulated peak imbalance compared with selected baselines. The evidence supports further controlled evaluation, not universal deployment. Component-matched ablations, calibrated intervals, multi-community tests, and field validation are needed before strong generalization claims can be made.

References

[1] A. Moreno and T. Li. Seasonal models for aggregated residential demand. Journal of Synthetic Energy Analytics, 2021.

[2] P. Ndlovu, E. Rossi, and M. Chen. Tree-based short-term forecasting with weather interactions. Proceedings of the Fictional Power Systems Conference, 2022.

[3] K. Ito and S. Rahman. Recurrent representations for neighborhood load prediction. Synthetic Transactions on Smart Energy, 2020.

[4] J. Bell, R. Mensah, and Y. Klein. Interpretable temporal fusion for multi-horizon energy forecasts. Journal of Demonstration Machine Learning, 2023.

[5] L. Anders and H. Qureshi. Quantile learning for probabilistic electricity demand. Fictional Energy Informatics, 2021.

[6] C. Martins, O. Adeyemi, and F. Wu. Calibration diagnostics for operational forecast intervals. Journal of Synthetic Forecasting, 2024.

[7] R. Gomez and A. Stein. Robust battery scheduling under bounded demand error. Demonstration Journal of Energy Storage, 2020.

[8] V. Rios and N. Patel. Probabilistic recurrent scheduling for community batteries. Fictional Applied Energy Systems, 2023.

[9] M. Weber, J. Park, and I. Salim. Distributionally robust reserve control in local energy systems. Synthetic IEEE Energy Letters, 2024.