Japanese Journal of Statistics and Data Science
https://doi.org/10.1007/s42081-026-00353-1

ORIGINAL PAPER

Advanced Theory and Applications in Time Series Analysis

Smart grid electricity demand forecasting using weather-based MIDAS and machine learning models: the case of New Zealand

Rogith Ramesh Babu[1] · Nuttanan Wichitaksorn[1] · Shu Su[1] · Clarissa Cortes Pires[2] · Edna Lu[2]

Received: 20 February 2026 / Revised: 12 May 2026 / Accepted: 20 May 2026

Abstract

This study develops a comparative forecasting framework that integrates daily weather information with quarterly electricity generation, used here as a proxy for electricity demand in New Zealand, through mixed-frequency modelling approaches. The analysis progresses from baseline univariate time-series models to classical mixed data sampling regressions, advanced regularised and autoregressive mixed-frequency models, and machine learning-based mixed-frequency methods. The forecasting results show that mixed-frequency models can improve upon traditional univariate benchmarks by incorporating higher-frequency weather information. Among the advanced approaches, autoregressive mixed-frequency models deliver strong forecasting performance, particularly over shorter recent evaluation windows, while seasonal time-series benchmarks such as SARIMA remain highly competitive and achieve the lowest RMSE in the main eight-quarter evaluation period. Machine learning-based mixed-frequency models show mixed performance, likely reflecting the challenges posed by data dimensionality and limited sample size. The proposed framework provides interpretable forecasts and offers practical insights for electricity system planning in renewable-dominated energy systems.

Keywords Electricity demand forecasting · Mixed-frequency modelling · Weather-based forecasting · Machine learning · New Zealand energy system

✉ Nuttanan Wichitaksorn
nuttanan.wichitaksorn@aut.ac.nz

1 Department of Mathematical Sciences, Auckland University of Technology, Auckland, New Zealand

2 Mercury Energy, Auckland, New Zealand

Published online: 02 June 2026

1 Introduction

Accurate electricity demand forecasting is a fundamental component of power system planning and operation. Reliable forecasts support generation scheduling, network investment decisions, congestion management, and the secure integration of renewable energy resources within smart grid environments. Electricity demand is inherently influenced by a range of exogenous drivers, among which weather-related factors such as temperature, seasonal patterns, and climatic variability play a dominant role. Numerous studies have documented the strong and often nonlinear relationship between electricity consumption and weather conditions, highlighting the importance of incorporating meteorological information into demand forecasting models (Bessec & Fouquau, 2008; Mirasgedis et al., 2006; Pardo et al., 2002).

Beyond electricity demand forecasting, a substantial body of literature has examined electricity price forecasting across short-, medium-, and long-term horizons, emphasising the importance of robust and flexible modelling frameworks for energy system planning and market operations (Ghelasi & Ziel, 2025; Kapoor et al., 2025; Nowotarski & Weron, 2018; Ziel & Weron, 2018). These studies highlight key characteristics of electricity systems, including strong temporal dependence, nonlinear dynamics, and sensitivity to exogenous drivers such as weather and market conditions. Although this study focuses on electricity demand rather than prices, insights from this literature reinforce the need for forecasting models that can effectively integrate higher-frequency information and accommodate complex temporal structures.

Traditional approaches to electricity demand forecasting have largely relied on classical time-series techniques, including ARIMA and SARIMA models, as well as exponential smoothing methods (Amjady, 2001; Gardner, 2006; Ho & Xie, 1998; Paul et al., 2018). These models remain widely adopted in practice due to their interpretability, relatively low computational cost, and solid theoretical foundations. Forecast accuracy is typically assessed using standard error-based measures and formal statistical comparison procedures, and commonly used accuracy metrics (Hyndman & Koehler, 2006). While effective as baseline models, traditional time-series approaches often face limitations when attempting to capture complex demand dynamics driven by higher frequency weather information, particularly when the target electricity demand series is observed at a lower temporal resolution, and that results in the mixed-frequency data for the analysis.

The challenge of incorporating higher-frequency explanatory variables into low-frequency forecasting models has motivated the development of mixed-frequency econometric frameworks. Mixed Data Sampling (MIDAS) regression models provide a flexible and statistically principled approach for exploiting information from higher-frequency covariates without the need for temporal aggregation (Clements & Galvão, 2008; Ghysels et al., 2007). By allowing daily or sub-daily weather observations to directly influence quarterly or monthly electricity demand, MIDAS models can preserve valuable information that is otherwise lost through averaging. Further extensions, such as unrestricted MIDAS formulations, relax parametric lag constraints and enhance model flexibility, enabling improved representation of complex lag structures (Foroni et al., 2015). Mixed-frequency models have demonstrated strong performance

in a range of macroeconomic, financial, and energy-related forecasting applications (Andreou et al., 2013; Ghysels et al., 2006).

In parallel with econometric developments, machine learning methods have gained increasing prominence in electricity demand forecasting. Algorithms such as random forests, gradient boosting machines, and related tree-based models have shown strong predictive performance by capturing nonlinear relationships and high-order interactions among explanatory variables (Breiman, 2001; Chen & Guestrin, 2016; Friedman, 2001). These methods have been successfully applied in short-term load forecasting and smart grid contexts (Ibrahim et al., 2022; Lahouar & Ben Hadj Slama, 2015; Magalhães, 2024). Despite their empirical success, many machine learning approaches rely on heuristic feature engineering and temporal aggregation when dealing with mixed-frequency data, which may reduce interpretability and obscure the underlying temporal structure of weather effects.

The New Zealand electricity system presents a particularly compelling setting for demand forecasting research. New Zealand has one of the highest shares of renewable electricity generation globally and has articulated a national ambition to transition towards a near 100% renewable electricity system by 2030 (Ministry of Business, 2019, 2023).

Electricity demand and generation in New Zealand are strongly influenced by weather variability due to the dominance of hydroelectric, wind, and geothermal resources. Note that, from an industry perspective, the real-time electricity generation equals demand. As the electricity system becomes increasingly renewable and weather-dependent, accurate medium-term demand forecasting becomes critical for system planning, market design, and operational resilience. Publicly available datasets from the Ministry of Business, Innovation and Employment and the National Institute of Water and Atmospheric Research provide a rich empirical basis for analysing the interaction between electricity demand and weather conditions in this context (Ministry of Business, 2024; National Institute of Water and Atmospheric Research, 2024).

Motivated by these challenges, this study examines electricity demand forecasting for New Zealand using a comparative modelling framework that integrates baseline time-series models, classical and advanced MIDAS regression approaches, and machine learning-based MIDAS extensions. The proposed framework seeks to leverage higher-frequency weather information while maintaining statistical structure and interpretability. By systematically comparing forecasting performance across model classes, this study provides empirical insights into the relative strengths of econometric and machine learning approaches for medium-term electricity demand forecasting in smart grid systems undergoing rapid energy transition.

The remaining structure of this paper is as follows: Sect. 2 describes the data and the methodology used. Section 3 presents the results. Section 4 discusses the implications and limitations. Section 5 concludes.

2 Data and Methodology

2.1 Data

This study integrates New Zealand electricity demand data from the Ministry of Business, Innovation and Employment (MBIE) with higher-frequency meteorological observations from the National Institute of Water and Atmospheric Research (NIWA). The two datasets differ substantially in temporal resolution, with electricity demand observed at a quarterly frequency and weather variables recorded daily. Such mixed-frequency data structures are common in energy and macroeconomic applications and pose challenges for conventional forecasting methods that rely on temporally aligned data. This motivates the use of mixed-frequency modelling frameworks, such as Mixed Data Sampling (MIDAS) regressions, which allow higher-frequency covariates to be incorporated without temporal aggregation (Clements & Galvão, 2008; Ghysels et al., 2007).

2.1.1 Electricity Demand

Quarterly electricity demand data were obtained from MBIE and consist of net electricity generation measured in gigawatt-hours (GWh). The dataset spans the period from 1990Q1 to 2025Q2, providing a long historical record suitable for medium-term electricity demand analysis. Net generation represents total electricity produced after accounting for station use and losses and is commonly used as a system-level proxy for electricity demand in empirical energy studies (Mirasgedis et al., 2006; Pardo et al., 2002). Note that, in this context, the net generation is treated as a practical system-level proxy for realised electricity consumption, which is demand, since publicly available quarterly load data are limited.

In addition to the aggregate demand series, the MBIE dataset includes electricity generation by major technology types, including hydro, geothermal, wind, solar, and bioenergy. These series capture important structural features of the New Zealand electricity system and are retained as potential explanatory variables for multivariable baseline models and mixed-frequency regressions. All calendar dates were converted to quarter-end timestamps to ensure consistent temporal alignment across variables.

Exploratory data analysis revealed a small number of missing observations in the wind generation series, while all other electricity generation variables were complete. To ensure comparability across modelling approaches and avoid unnecessary loss of information, the electricity demand series and all selected quarterly covariates were enforced to contain no missing values. Missing quarterly observations were addressed using time-based interpolation, followed by forward and backward filling at the sample boundaries where necessary. This approach is widely used in energy demand studies when missingness is limited and avoids introducing artificial discontinuities in the data (Little & Rubin, 2019).

Seasonal analysis of quarterly electricity demand reveals a clear and persistent seasonal pattern, with systematic differences across calendar quarters. Such seasonality is consistent with previous findings in the electricity demand literature and reflects the influence of temperature variation, hydrological conditions, and seasonal heating

demand (Bessec & Fouquau, 2008; Hahn et al., 2009). These features motivate the inclusion of both seasonal components and weather-related predictors in subsequent modelling stages.

2.1.2 Weather Data

Daily weather data were sourced from NIWA and include observations from multiple meteorological stations distributed across New Zealand. The weather variables considered in this study are maximum temperature, minimum temperature, mean temperature, rainfall, and wind speed. The dataset covers the same calendar period as the electricity demand data, from January 1990 to June 2025, allowing consistent alignment across modelling stages.

Weather variables exhibit strong seasonal patterns, substantial short-term variability, and occasional extreme values, particularly for rainfall and wind speed. Such behaviour is characteristic of New Zealand's climate and has been documented extensively in the energy-weather literature (Bessec & Fouquau, 2008; Taylor & Buizza, 2003). Rather than removing extreme observations, a conservative winsorisation procedure was applied to all weather variables at the 0.5th and 99.5th percentiles. This approach limits the influence of extreme spikes while preserving the underlying distributional properties of the data and is commonly adopted in empirical forecasting studies (Box & Cox, 1964).

Missing observations are common in station-level weather datasets due to measurement interruptions and reporting gaps. To minimise distortion of temporal dynamics, only short missing gaps of up to seven consecutive days were interpolated using time-based interpolation. Longer gaps were intentionally left unfilled. This strategy balances data completeness with the need to preserve realistic variability in higher-frequency weather series and avoids excessive imputation that could bias model estimation (Little & Rubin, 2019). After this procedure, a non-trivial proportion of missing observations remained for certain stations, reflecting prolonged data gaps that were deliberately not imputed.

No scaling, normalisation, or temporal aggregation of weather variables was performed during preprocessing. All transformations related to scaling are conducted within individual modelling pipelines to prevent information leakage between training and evaluation samples, in line with best practices in forecasting and machine learning studies (Bontempi et al., 2013; Hyndman & Koehler, 2006).

2.1.3 Mixed-Frequency Alignment

The final prepared datasets preserve the original temporal resolution of each data source. Electricity demand remains observed at the quarterly frequency, while weather variables remain at the daily frequency. No temporal aggregation of weather data was applied during preprocessing. This design choice allows mixed-frequency models to exploit the full information content of higher-frequency weather observations without loss due to averaging, which is a key advantage of MIDAS-type frameworks (Foroni et al., 2015; Ghysels et al., 2007).

Two clean and internally consistent datasets were produced as outputs of the preprocessing stage: a quarterly MBIE dataset containing electricity demand and selected generation components, and a daily NIWA dataset containing weather variables across all stations. These datasets form the empirical foundation for all subsequent baseline time-series models, mixed-frequency MIDAS regressions, and machine learning-based MIDAS extensions developed in this study.

2.2 Methodology: Model and Methods

This section outlines the forecasting framework used to model quarterly electricity demand in New Zealand using both time-series and mixed-frequency approaches. The methodology is structured in a progressive manner, beginning with simple benchmark models and gradually incorporating higher model complexity and additional information. This hierarchical design allows the incremental value of mixed-frequency weather information and advanced modelling techniques to be assessed in a transparent and systematic way (Hyndman & Koehler, 2006; Shumway & Stoffer, 2017).

First, a set of baseline time-series forecasting models is employed to establish reference forecasts based solely on historical electricity demand. These benchmarks are commonly used in empirical forecasting studies and provide a necessary baseline against which more complex approaches can be evaluated (Gardner, 2006; Ho & Xie, 1998; Paul et al., 2018). Next, classical mixed data sampling (MIDAS) regression models are introduced to incorporate higher-frequency weather information through lag weighting structures without temporal aggregation, which is a key advantage in mixed-frequency forecasting settings (Clements & Galvão, 2008; Foroni et al., 2015; Ghysels et al., 2007).

Building on this framework, advanced MIDAS specifications are considered, including regularised extensions that reduce overfitting in high-dimensional lag settings and autoregressive mixed-frequency formulations that account for persistence in demand dynamics (Hoerl & Kennard, 1970; Tibshirani, 1996; Zou & Hastie, 2005). Finally, machine learning-based MIDAS models are implemented to capture nonlinear relationships and complex interactions between demand and higher-frequency weather predictors, which are frequently observed in smart grid and load forecasting applications (Bontempi et al., 2013; Breiman, 2001; Chen & Guestrin, 2016; Friedman, 2001).

All models are estimated and evaluated using a consistent expanding-window forecasting design to reflect realistic real-time forecasting conditions and to ensure fair model comparison. Forecast accuracy is assessed using established forecast error measures, i.e., root mean squared errors, mean absolute errors, mean absolute percentage errors, and mean absolute scaled error (Hyndman & Koehler, 2006).

2.2.1 Baseline Forecasting Models

Baseline forecasting models are employed to establish benchmark performance levels based solely on historical electricity demand. These models do not incorporate exogenous information and serve as reference points against which more sophisti-

cated mixed-frequency approaches can be evaluated. Benchmark models are widely recommended in forecasting studies to ensure that gains from increased model complexity are meaningful and robust (Gardner, 2006; Hyndman & Koehler, 2006). The benchmark models include naïve, seasonal naïve, ARIMA, seasonal ARIMA, and exponential smoothing state-space models.

2.2.2 Classical MIDAS Regression Models

In this study, Mixed Data Sampling (MIDAS) regression models are employed to link daily meteorological variables from NIWA to quarterly electricity demand in New Zealand. See (Clements & Galvão, 2008; Ghysels et al., 2007) for more details on MIDAS. This approach allows weather conditions leading up to each quarter to influence demand forecasts through flexible lag-weighting schemes, while preserving the original data frequencies. This is particularly suitable for smart grid and energy system applications, where short-term weather dynamics play a critical role in shaping electricity demand patterns (Bontempi et al., 2013; Foroni et al., 2015).

Let y_t denote quarterly electricity demand observed at quarter t, and let $x_{t,d}^{(k)}$ represent the k-th daily weather variable observed on day d prior to quarter t. A general MIDAS regression model with multiple higher-frequency predictors can be written as:

$$y_t = \beta_0 + \sum_{k=1}^{K} \beta_k \sum_{h=1}^{H} w_h^{(k)}(\boldsymbol{\theta}_k)\, x_{t-h}^{(k)} + \varepsilon_t,$$

where β_0 is an intercept term, β_k are slope coefficients associated with each predictor, $w_h^{(k)}(\cdot)$ are lag weights parameterised by $\boldsymbol{\theta}_k$, H denotes the maximum number of daily lags included, and ε_t is an error term with zero mean, σ standard deviation, and normally distributed. The set of higher-frequency predictors includes daily mean temperature, rainfall, and wind speed measured across multiple meteorological stations. Each weather variable contributes to quarterly demand through a weighted average of its recent daily observations, allowing the model to capture delayed and distributed weather effects.

To parsimoniously model the lag structure of higher-frequency predictors, MIDAS regressions commonly employ parametric lag polynomials. Following Ghysels et al. (2007), the Beta lag polynomial is adopted in this study due to its flexibility and low dimensionality.

The Beta weighting function is defined as:

$$w_h(a,b) = \frac{\left(\frac{h}{H}\right)^{a-1}\left(1-\frac{h}{H}\right)^{b-1}}{\sum_{j=1}^{H}\left(\frac{j}{H}\right)^{a-1}\left(1-\frac{j}{H}\right)^{b-1}}, \quad h = 1, \ldots, H,$$

where $a > 0$ and $b > 0$ are shape parameters controlling the decay and concentration of lag weights. This formulation allows the influence of past weather observations to peak early, late, or in the middle of the lag window, depending on the estimated parameters. Separate Beta lag parameters are estimated for temperature, rainfall, and

wind speed, enabling each weather variable to exhibit distinct temporal impact patterns on electricity demand.

The Beta lag weighting function can also be interpreted in terms of the memory effect of higher-frequency weather variables on the quarterly target. In general, the estimated shape parameters (a, b) determine whether the implied weighting pattern places more emphasis on recent observations, more distant lags, or a more distributed set of lags across the weighting window. Greater weight on recent lags suggests a shorter memory effect, meaning that more recent weather conditions are more informative for forecasting. In contrast, a more distributed weighting pattern suggests that the effect of weather persists over a longer part of the lag window.

The choice of lag length H determines the temporal window over which daily weather observations influence quarterly electricity demand. Rather than fixing H a priori, we evaluate multiple lag lengths to assess the sensitivity of MIDAS forecasts to the assumed memory of weather effects. Specifically, lag lengths of $H = 90$, 120, and 150 days are considered. These choices reflect plausible physical relationships between weather conditions and electricity demand, such as cumulative heating or cooling effects and hydrological persistence.

For each lag length, the MIDAS model is estimated and evaluated using an identical expanding-window forecasting scheme. Comparing forecast accuracy across different values of H allows the most informative lag horizon to be identified empirically, while maintaining consistency across model specifications (Clements & Galvão, 2008; Foroni et al., 2015).

The MIDAS regression model is estimated using an expanding-window forecasting design to replicate real-time prediction conditions. At each forecast origin, all available historical quarterly demand observations and corresponding daily weather data are used to estimate model parameters, after which a one-step-ahead forecast is produced for the next quarter.

To improve numerical stability and mitigate potential multicollinearity arising from multiple higher-frequency predictors, ridge regression is employed when estimating the MIDAS regression coefficients. The ridge estimator solves:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \left\{ \sum_{t=1}^{T} (y_t - \mathbf{X}_t \boldsymbol{\beta})^2 + \lambda \sum_{j=1}^{J} \beta_j^2 \right\},$$

where λ is a small regularisation parameter and the intercept term is excluded from penalisation.

All preprocessing steps, including standardisation of daily weather variables, are performed using training data only at each forecasting step to prevent information leakage. Forecast accuracy is evaluated over the final eight quarters of the sample.

Classical MIDAS regressions provide a flexible way to incorporate higher-frequency weather information into low-frequency electricity demand forecasts, but their performance can deteriorate when many predictors are included or when the target series exhibits strong persistence. The use of multiple meteorological stations and weather variables can create high-dimensional regressor spaces, increasing the risks of multicollinearity and overfitting, while standard MIDAS specifications do not explic-

itly account for the pronounced temporal dependence in electricity demand. To address these issues, advanced MIDAS extensions are considered, namely regularised MIDAS regressions and autoregressive MIDAS models. Regularisation introduces shrinkage to stabilise estimation in high-dimensional settings (Hoerl & Kennard, 1970; Tibshirani, 1996; Zou & Hastie, 2005), whereas autoregressive MIDAS formulations incorporate lagged demand to capture persistence alongside higher-frequency weather effects (Foroni et al., 2015). Together, these approaches enhance model stability, dynamic structure, and forecasting robustness, which is particularly relevant for New Zealand's electricity demand given its strong seasonality and sensitivity to weather-dependent renewable generation.

2.2.3 Ridge-MIDAS (Regularised MIDAS Regression)

Regularisation is a standard solution in such settings, where a penalty is introduced to stabilise coefficient estimation and reduce overfitting (Hoerl & Kennard, 1970). In this study, an advanced MIDAS specification referred to as Ridge-MIDAS is implemented by combining a beta-weighted MIDAS feature construction with ridge-regularised regression. For each quarter t, the MIDAS transformation builds a weighted summary of the previous H daily observations using a beta weighting scheme (Ghysels et al., 2007):

$$z_t^{(j)}(H, \theta_j) = \sum_{k=1}^{H} w_k(H, \theta_j)\, x_{t-k+1}^{(j)}, \tag{1}$$

where $w_k(H, \theta_j)$ are normalised beta weights (reversed so that higher weights can be assigned to more recent days, depending on the estimated shape). In this work, the predictors are grouped into three families, mean temperature, rainfall, and wind speed, and a separate beta shape parameter pair is estimated for each family:

$$\theta_{\text{temp}} = (a_{\text{temp}}, b_{\text{temp}}), \quad \theta_{\text{rain}} = (a_{\text{rain}}, b_{\text{rain}}), \quad \theta_{\text{wind}} = (a_{\text{wind}}, b_{\text{wind}}). \tag{2}$$

This design matches the empirical setting where multiple stations contribute daily observations for each weather variable family.

After the MIDAS features are constructed, quarterly demand is modelled using a linear regression with ridge regularisation. Let $\mathbf{z}_t \in \mathbb{R}^p$ be the vector of all MIDAS-transformed predictors for quarter t, where p is the number of MIDAS-transformed weather features obtained from the selected stations and variable families.

The Ridge-MIDAS regression is:

$$y_t = \beta_0 + \mathbf{z}_t^\top \boldsymbol{\beta} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2),$$

where coefficients are estimated by minimising the penalised sum of squares:

$$\min_{\beta_0, \boldsymbol{\beta}} \sum_{t \in \mathcal{T}_{\text{train}}} \left(y_t - \beta_0 - \mathbf{z}_t^\top \boldsymbol{\beta} \right)^2 + \lambda \|\boldsymbol{\beta}\|_2^2. \tag{3}$$

The intercept β_0 is not penalised, consistent with standard ridge regression practice (Hoerl & Kennard, 1970).

Following the baseline MIDAS results, the horizon was fixed at $H = 120$ days because it provided the best performance among the candidate horizons tested. For each expanding-window fold, the beta-kernel parameters in (1)–(2) are estimated by numerical optimisation (Nelder-Mead) to minimise the training sum of squared errors, using a small stabilising ridge penalty during optimisation. Once the MIDAS features are fixed for that fold, the ridge penalty λ in (3) is selected using a simple grid search over candidate values:

$$\lambda \in \{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1.0\}.$$

This procedure keeps the estimation stable while remaining computationally feasible under a rolling forecasting setup. Then, the Ridge-MIDAS is evaluated using the same expanding-window forecasting design adopted across all models. For reproducibility, each fold produces the one-step-ahead forecast $\hat{y}_t$ and stores the selected ridge penalty λ and the estimated beta shape parameters for each weather family. The corresponding prediction table and the plot comparing Ridge-MIDAS forecasts against actual values for the last eight quarters are reported in the Results section.

2.2.4 AR-MIDAS (Autoregressive MIDAS Regression)

The autoregressive (AR)-MIDAS model augments the MIDAS regression with an AR(1) term:

$$y_t = \beta_0 + \phi\, y_{t-1} + \mathbf{z}_t^{\top}\boldsymbol{\beta} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2),$$

where ϕ captures quarterly persistence and $\mathbf{z}_t \in \mathbb{R}^p$ is the stacked vector of MIDAS-transformed daily predictors across all selected stations, where p denotes the total number of MIDAS-transformed weather features. The modelling and estimation procedures follow as those in the Ridge-MIDAS mentioned above.

2.2.5 Machine learning MIDAS models

In addition to the classical and advanced MIDAS regressions, this study evaluates a set of machine learning MIDAS (ML-MIDAS) models to capture potentially nonlinear relationships between daily weather conditions and quarterly net electricity generation. Mixed-frequency forecasting is well motivated in macroeconomic and energy applications because higher-frequency information, such as daily weather observations, can be informative for lower-frequency targets, while conventional single-frequency models cannot directly exploit this structure (Clements & Galvão, 2008; Ghysels et al., 2007). Electricity demand and generation are known to respond to weather conditions through seasonality and persistence (Mirasgedis et al., 2006; Pardo et al., 2002), as well as through nonlinear effects (Bessec & Fouquau, 2008; Hong & Fan, 2016), motivating the use of flexible forecasting methods.

The ML-MIDAS framework follows a two-stage approach. First, daily NIWA weather variables (mean temperature, rainfall, and wind speed) are aggregated into quarterly MIDAS features using a fixed Beta lag-weighting scheme. Second, machine learning models are trained on the resulting quarterly feature matrix to predict quarterly net electricity generation. This design preserves the MIDAS principle of summarising higher-frequency predictors into low-frequency regressors (Foroni et al., 2015; Ghysels et al., 2007), while allowing a flexible, potentially nonlinear mapping from predictors to the target variable via machine learning methods (Bontempi et al., 2013; Ibrahim et al., 2022).

The ML-MIDAS models used in this study were chosen to provide a broader comparison within the mixed-frequency forecasting framework. While classical and advanced MIDAS regressions offer a structured way to link daily weather information with quarterly net electricity generation, they are mainly designed around linear relationships. In practice, the effect of weather on net electricity generation may be more complex, especially when interactions, threshold effects, or changing seasonal patterns are present. For this reason, the ML-MIDAS framework is used to examine whether machine learning models can provide additional forecasting improvements once the higher-frequency weather variables have been transformed into quarterly MIDAS features.

The models considered in this study represent two main groups. The first group includes LASSO-MIDAS and ElasticNet-MIDAS, which are regularised linear models. These models are useful when there are many correlated predictors because they can shrink less important coefficients, improve estimation stability, and produce more interpretable results. The second group includes RF-MIDAS, GBR-MIDAS, and XGB-MIDAS, which are tree-based models. These models are more flexible and can capture nonlinear patterns, interactions between predictors, and changes in behaviour under different weather conditions. By considering both groups, the ML-MIDAS framework allows a direct comparison between simpler linear models and more flexible nonlinear models using the same mixed-frequency weather features.

Given the constructed MIDAS feature matrix, the ML-MIDAS forecasting problem is expressed as:

$$y_t \;=\; f(\mathbf{z}_t) + \varepsilon_t,$$

where $f(\cdot)$ is a machine learning regression function, which may be either linear or nonlinear depending on the selected model. For example, in LASSO-MIDAS and ElasticNet-MIDAS, $f(\mathbf{z}_t)$ takes a linear form such as

$$f(\mathbf{z}_t) = \beta_0 + \mathbf{z}_t^\top \boldsymbol{\beta}.$$

In contrast, for RF-MIDAS, GBR-MIDAS, and XGB-MIDAS, $f(\mathbf{z}_t)$ is a nonlinear function represented by an ensemble of regression trees. This formulation allows the ML-MIDAS framework to include both simpler linear models and more flexible nonlinear models within the same forecasting setting.

Five ML-MIDAS models are considered in this study: LASSO-MIDAS and ElasticNet-MIDAS as regularised linear benchmarks, and RF-MIDAS, GBR-MIDAS, and XGB-MIDAS as nonlinear tree-based learning models (Breiman, 2001; Chen &

Guestrin, 2016; Friedman, 2001; Tibshirani, 1996; Zou & Hastie, 2005). Together, these models allow the forecasting performance of sparse linear methods and more flexible nonlinear methods to be compared within the same ML-MIDAS setting.

All ML-MIDAS models are evaluated using a rolling-origin forecasting scheme over the final eight quarters of the sample. Within each fold, all scaling operations are performed using training data only to avoid information leakage. Daily predictors are standardised using the training-period mean and standard deviation prior to MIDAS aggregation, and for linear models (LASSO and Elastic Net), the quarterly MIDAS feature matrix is additionally standardised using training data only. Hyperparameters for LASSO and Elastic Net are selected using an inner validation split within each training window.

LASSO-MIDAS The first machine learning MIDAS specification considered is the LASSO-MIDAS model, which combines MIDAS-based mixed-frequency aggregation with ℓ_1-penalised linear regression. LASSO-MIDAS is well suited to high-dimensional MIDAS settings where the number of aggregated predictors may be large and potentially correlated, as it performs both coefficient shrinkage and automatic variable selection (Tibshirani, 1996). This property is particularly relevant in electricity forecasting applications where multiple weather stations and meteorological variables are incorporated.

Let $\mathbf{z}_t = (z_t^{(1)}, \ldots, z_t^{(p)})^\top$ denote the vector of MIDAS-aggregated weather predictors at quarter t. The LASSO-MIDAS regression model is defined as:

$$y_t = \beta_0 + \sum_{j=1}^{p} \beta_j z_t^{(j)} + \varepsilon_t,$$

where y_t is quarterly net electricity generation and ε_t is an error term with zero mean, σ standard deviation, and normally distributed. Parameter estimation is obtained by solving the ℓ_1-penalised least squares problem:

$$\widehat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \left\{ \frac{1}{T} \sum_{t=1}^{T} \left(y_t - \beta_0 - \sum_{j=1}^{p} \beta_j z_t^{(j)} \right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\},$$

where $\lambda \geq 0$ controls the degree of shrinkage and sparsity (Tibshirani, 1996). Larger values of λ result in stronger regularisation and more coefficients being shrunk exactly to zero.

To ensure comparability across predictors and numerical stability, the MIDAS feature matrix $\mathbf{z}_t$ is standardised using training-sample means and standard deviations prior to estimation. This step is essential for ℓ_1-regularised models, as the penalty term depends on the scale of the regressors (Zou & Hastie, 2005). The intercept is not penalised. The regularisation parameter λ is selected within each rolling training window using an inner validation split of the quarterly data. For each candidate value of λ, the model is fitted on the inner training subset and evaluated on the validation subset using root mean squared error. The value of λ minimising validation error is

retained and used to estimate the final LASSO-MIDAS model for that forecasting origin.

Within the ML-MIDAS framework, LASSO-MIDAS serves as a parsimonious linear benchmark that extends classical MIDAS regressions by allowing automatic predictor selection across a large set of aggregated weather variables. This makes it a useful reference model for assessing the incremental value of more flexible non-linear ML-MIDAS approaches introduced in subsequent sections.

ElasticNet-MIDAS The ElasticNet-MIDAS model extends LASSO-MIDAS by combining ℓ_1 and ℓ_2 regularisation within the MIDAS-based mixed-frequency framework. While LASSO-MIDAS performs effective variable selection, it can become unstable when predictors are highly correlated, a situation that commonly arises when multiple weather variables and stations are included. Elastic Net regularisation addresses this limitation by encouraging grouped selection of correlated predictors while retaining sparsity (Zou & Hastie, 2005).

Following the same setup as LASSO-MIDAS, the ElasticNet-MIDAS regression model is defined as:

$$y_t \;=\; \beta_0 + \sum_{j=1}^{p} \beta_j z_t^{(j)} + \varepsilon_t,$$

where y_t denotes quarterly net electricity generation and ε_t is a zero-mean disturbance term. Parameter estimation is obtained by solving the Elastic Net penalised least squares problem:

$$\widehat{\boldsymbol{\beta}} \;=\; \arg\min_{\boldsymbol{\beta}} \left\{ \frac{1}{T} \sum_{t=1}^{T} \left(y_t - \beta_0 - \sum_{j=1}^{p} \beta_j z_t^{(j)} \right)^2 \;+\; \lambda \left[\alpha \sum_{j=1}^{p} |\beta_j| + (1-\alpha) \sum_{j=1}^{p} \beta_j^2 \right] \right\},$$

where $\lambda \geq 0$ controls the overall strength of regularisation and $\alpha \in [0, 1]$ determines the balance between ℓ_1 and ℓ_2 penalties (Zou & Hastie, 2005). Setting $\alpha = 1$ recovers the LASSO-MIDAS model, while $\alpha = 0$ corresponds to ridge-type shrinkage (Hoerl & Kennard, 1970).

As with LASSO-MIDAS, the MIDAS feature matrix is standardised using training-sample means and standard deviations prior to estimation to ensure scale invariance of the penalty terms. The intercept is excluded from penalisation. Daily weather variables are standardised prior to MIDAS aggregation to avoid information leakage across forecasting origins. However, the ElasticNet-MIDAS model requires selection of two hyperparameters, λ and α. Within each rolling training window, these parameters are selected using an inner validation split of the quarterly data. Candidate values of (λ, α) are evaluated based on validation root mean squared error, and the optimal combination is used to refit the model on the full training sample for out-of-sample forecasting.

ElasticNet-MIDAS provides a compromise between sparsity and stability in high-dimensional MIDAS regressions. By allowing correlated weather predictors to enter the model jointly, it offers a more robust linear benchmark than LASSO-MIDAS while remaining interpretable. This makes ElasticNet-MIDAS a useful reference point for

evaluating the gains from non-linear tree-based ML-MIDAS models considered in subsequent sections.

Random Forest-MIDAS (RF-MIDAS) The Random Forest-MIDAS (RF-MIDAS) model extends the mixed-frequency forecasting framework by allowing for non-linear and interaction effects between MIDAS-aggregated weather variables. Electricity generation and demand are known to respond to weather conditions in a complex and non-linear manner, particularly at seasonal extremes, which can limit the performance of linear MIDAS regressions (Bessec & Fouquau, 2008; Hong & Fan, 2016; Mirasgedis et al., 2006). Tree-based ensemble methods provide a flexible alternative that can automatically capture such effects without explicit model specification.

The RF-MIDAS model assumes the following non-linear regression structure:

$$y_t \;=\; f_{\mathrm{RF}}(\mathbf{z}_t) + \varepsilon_t,$$

where $f_{\mathrm{RF}}(\cdot)$ is an ensemble of regression trees and ε_t is a zero-mean error term.

Random Forest regression approximates $f_{\mathrm{RF}}(\cdot)$ by averaging the predictions of B decision trees:

$$\widehat{f}_{\mathrm{RF}}(\mathbf{z}) \;=\; \frac{1}{B}\sum_{b=1}^{B} T_b(\mathbf{z}),$$

where each tree $T_b(\cdot)$ is trained on a bootstrap sample of the training data and constructed using recursive binary splits on randomly selected subsets of predictors (Breiman, 2001). This combination of bootstrap aggregation and random feature selection reduces variance and improves out-of-sample stability relative to single decision trees.

Within each rolling training window, RF-MIDAS is estimated using the MIDAS feature matrix without additional linear scaling, as tree-based models are invariant to monotonic transformations of the inputs. Key hyperparameters include the number of trees B, maximum tree depth, and minimum node size, which jointly control the bias-variance trade-off of the ensemble (Breiman, 2001). These hyperparameters are selected using an inner validation split within the training sample to minimise validation root mean squared error. For each forecasting origin, the RF-MIDAS model is trained on all available historical quarterly observations, and a one-step-ahead forecast is generated for the next quarter. This rolling-origin evaluation reflects real-time forecasting conditions and ensures a fair comparison with linear MIDAS and baseline time-series models.

Gradient Boosting Regression MIDAS (GBR-MIDAS) The Gradient Boosting Regression MIDAS (GBR-MIDAS) model further extends the ML-MIDAS framework by employing boosting-based tree ensembles to model potentially complex and non-linear relationships between MIDAS-aggregated weather variables and quarterly net electricity generation. Unlike bagging-based approaches such as Random Forests, boosting constructs the predictive model sequentially, allowing later learners to focus on correcting errors made by earlier ones (Friedman, 2001). This property is particularly attractive for electricity generation forecasting, where weather impacts may be

asymmetric and vary across operating regimes (Hong & Fan, 2016; Mirasgedis et al., 2006).

The GBR-MIDAS model assumes the additive structure

$$y_t \;=\; f_{\mathrm{GBR}}(\mathbf{z}_t) + \varepsilon_t,$$

where $f_{\mathrm{GBR}}(\cdot)$ is an additive ensemble of regression trees and ε_t is an error term with zero mean.

The boosting function $f_{\mathrm{GBR}}(\cdot)$ is constructed iteratively as

$$f_{\mathrm{GBR}}^{(M)}(\mathbf{z}) \;=\; \sum_{m=1}^{M} \nu\, T_m(\mathbf{z}),$$

where $T_m(\cdot)$ denotes the m-th regression tree, M is the total number of boosting iterations, and $\nu \in (0, 1]$ is a learning rate controlling the contribution of each tree (Friedman, 2001). Each tree is fitted to the negative gradient of the loss function with respect to the current model estimate, thereby sequentially reducing prediction errors.

In this study, squared error loss is used, consistent with standard regression boosting:

$$L(y_t, \widehat{y}_t) \;=\; (y_t - \widehat{y}_t)^2.$$

At each boosting iteration, the regression tree is fitted to the residuals from the previous stage, allowing the model to capture non-linearities and interaction effects between MIDAS features without explicit specification (Friedman, 2001).

GBR-MIDAS is estimated within each rolling training window using the MIDAS feature matrix. As with Random Forests, no additional feature scaling is required due to the tree-based structure. Key hyperparameters include the number of boosting iterations M, tree depth, and learning rate ν, which together govern the bias-variance trade-off of the ensemble. These parameters are selected using an inner validation split within the training sample to minimise validation forecast error.

Extreme Gradient Boosting MIDAS (XGB-MIDAS) The Extreme Gradient Boosting MIDAS (XGB-MIDAS) model represents the most flexible machine learning approach considered in this study. It combines the MIDAS aggregation framework with extreme gradient boosting, which extends classical gradient boosting by incorporating regularisation, shrinkage, and efficient optimisation strategies to improve predictive performance and generalisation (Chen & Guestrin, 2016). These properties are particularly relevant for electricity generation forecasting, where predictor dimensionality is high and relationships between weather variables and generation may be highly non-linear (Bontempi et al., 2013; Hong & Fan, 2016).

The XGB-MIDAS model also assumes the additive structure

$$y_t \;=\; f_{\mathrm{XGB}}(\mathbf{z}_t) + \varepsilon_t,$$

where $f_{\mathrm{XGB}}(\cdot)$ is an ensemble of regression trees learned via extreme gradient boosting and ε_t denotes a zero-mean error term.

The prediction function is expressed as a sum of M regression trees:

$$f_{\mathrm{XGB}}(\mathbf{z}) \;=\; \sum_{m=1}^{M} T_m(\mathbf{z}),$$

where each $T_m(\cdot)$ belongs to the space of decision trees. Trees are added sequentially, with each new tree fitted to reduce the residual errors of the existing ensemble (Chen & Guestrin, 2016).

Unlike standard gradient boosting, XGBoost explicitly regularises model complexity. The optimisation problem at iteration m minimises the regularised objective

$$\mathcal{L}^{(m)} \;=\; \sum_{t} \ell\Big(y_t, \widehat{y}_t^{(m-1)} + T_m(\mathbf{z}_t)\Big) \;+\; \Omega(T_m),$$

where $\ell(\cdot)$ denotes the squared error loss and $\Omega(T_m)$ is a complexity penalty on the tree structure:

$$\Omega(T) \;=\; \gamma|\mathcal{L}_T| + \frac{\lambda}{2}\sum_{j\in\mathcal{L}_T} w_j^2,$$

with $|\mathcal{L}_T|$ denoting the number of leaves in tree T, w_j the prediction weight of leaf j, and $\gamma, \lambda \geq 0$ regularisation parameters controlling model complexity (Chen & Guestrin, 2016).

The regularised boosting framework allows XGB-MIDAS to handle a large number of MIDAS features while mitigating overfitting. This is particularly important in mixed-frequency settings, where many daily predictors are aggregated into quarterly features and interactions between weather variables may be complex and non-linear (Hong & Fan, 2016; Mirasgedis et al., 2006). In contrast to RF-MIDAS and GBR-MIDAS, XGB-MIDAS jointly optimises fit and complexity at each boosting step, leading to more stable forecasts in high-dimensional environments (Bontempi et al., 2013).

Within the broader ML-MIDAS family, XGB-MIDAS serves as a high-capacity non-linear benchmark that balances flexibility and regularisation. Its inclusion allows assessment of whether highly structured boosting models provide additional forecasting gains over classical MIDAS regressions and simpler machine learning extensions in the context of mixed-frequency electricity generation forecasting (Chen & Guestrin, 2016; Hong & Fan, 2016).

3 Results

This section presents the empirical forecasting results for quarterly net electricity generation in New Zealand. Model performance is evaluated using a consistent rolling-origin forecasting design over the final eight quarters of the sample, ensuring comparability across all approaches. The results are organised progressively, beginning with exploratory analysis and baseline time-series benchmarks, followed by classical

MIDAS regressions, advanced MIDAS specifications, and finally machine learning-based MIDAS models.

3.1 Exploratory data analysis and preprocessing results

The results of the exploratory data analysis (EDA) and preprocessing applied to the MBIE quarterly electricity generation data and the NIWA daily weather dataset are illustrated to assess data quality, identify key structural patterns, and verify that the prepared datasets are suitable for subsequent baseline, mixed-frequency, and machine learning forecasting models.

3.1.1 MBIE electricity generation data

The MBIE dataset contains quarterly observations of net electricity generation and generation by major technology types over the period 1990Q1-2025Q2. Net electricity generation is used in this study as a system-level measure of realised electricity consumption. In New Zealand, supply and demand are continuously balanced, so total generation closely reflects actual electricity use over the same period.

The generation by technology series (hydro, geothermal, wind, solar, etc.) are included to describe the structure of the electricity system and its strong dependence on weather conditions. The sum of generation across technologies does not always exactly equal total net generation due to reporting adjustments, station use, network losses, and rounding in the official MBIE statistics. These small differences are common in national energy datasets and do not affect the modelling framework.

Initial inspection confirmed a complete and continuous quarterly time index with no duplicate observations.

Figure 1 presents boxplots of the raw MBIE variables. Net electricity generation (in GWh) exhibits a wide interquartile range reflecting long-term growth and variability in demand, while technology-specific generation series display heterogeneous dispersion patterns. Wind generation shows mild irregularities relative to other technologies, motivating closer inspection during preprocessing.

To illustrate recent demand dynamics, Fig. 2 plots net electricity generation over the most recent 10-year period. A clear cyclical pattern is observed, consistent with seasonal demand fluctuations, alongside moderate year-to-year variation.

Seasonal structure is further confirmed in Fig. 3, which shows the distribution of quarterly net generation by calendar quarter. The third quarter consistently records higher median generation levels, while the first quarter exhibits lower values, highlighting strong intra-year seasonality.

The missing values in the MBIE variables before and after preprocessing. Minor missingness was present only in the wind generation series and was fully resolved using time-based interpolation and boundary filling. All quarterly variables used in subsequent models are complete after preprocessing.

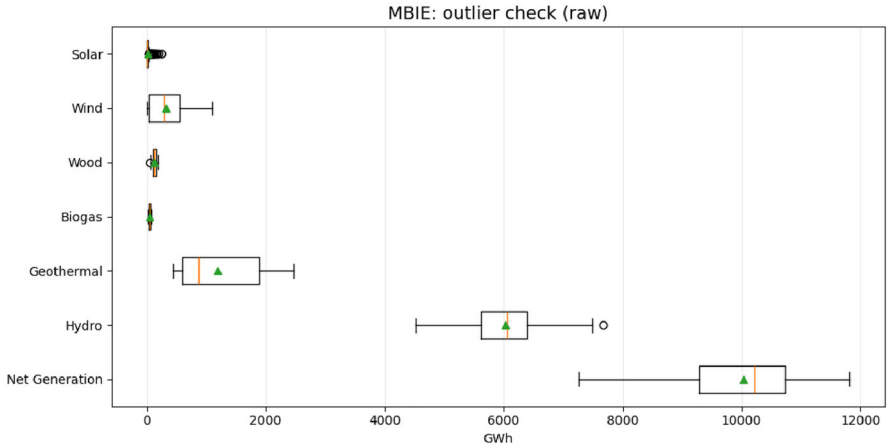


Fig. 1 Boxplots of raw MBIE quarterly variables illustrating variability and potential outliers across generation technologies

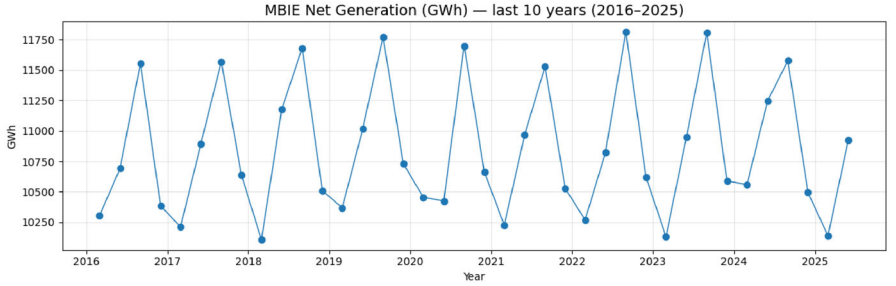


Fig. 2 MBIE quarterly net electricity generation over the last ten years

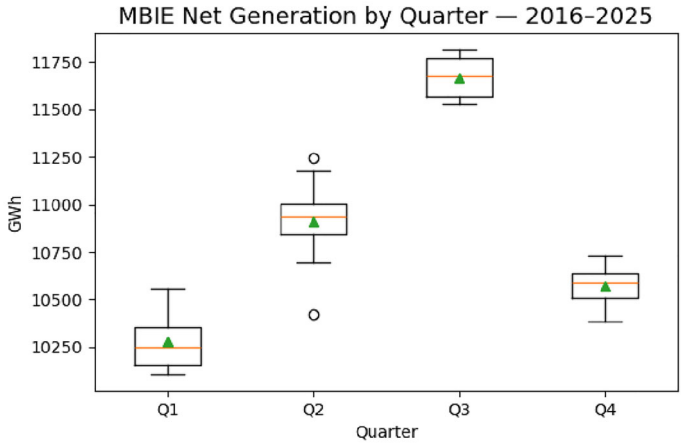


Fig. 3 Seasonal distribution of MBIE net electricity generation by quarter

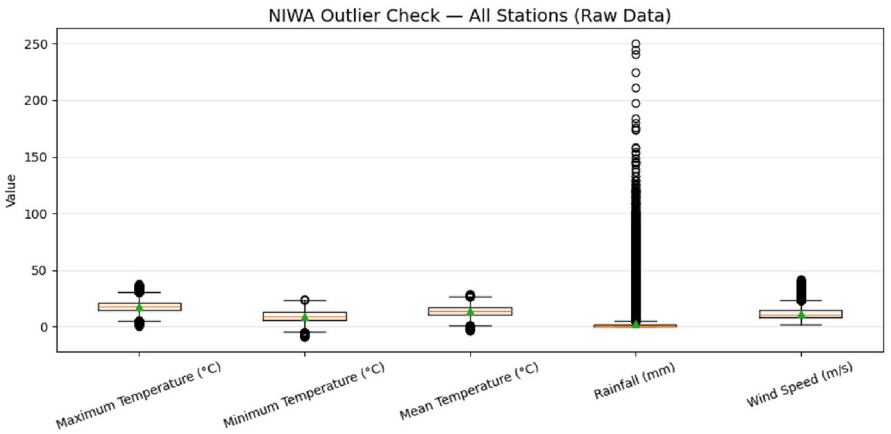


Fig. 4 Boxplots of NIWA daily weather variables across all stations (raw data)

3.1.2 NIWA daily weather data

The NIWA dataset comprises daily observations of temperature, rainfall, and wind speed from multiple meteorological stations across New Zealand over the same calendar period. Given the long temporal span and station-level granularity, missing observations are expected.

Figure 4 presents combined boxplots for all stations and weather variables. Rainfall exhibits pronounced right skewness with occasional extreme values, while temperature and wind speed show comparatively stable distributions. These characteristics are consistent with known climatic behaviour and motivate the application of conservative winsorisation rather than aggressive outlier removal.

Figure 5 illustrates daily weather dynamics for selected representative stations over the last ten years. Strong annual seasonality is evident in temperature series, while rainfall and wind display higher short-term variability and intermittent spikes.

After preprocessing, which included interpolation of short gaps only and winsorisation at the distribution tails, residual missingness remains for certain station-variable combinations. Table 1 reports the ten variables with the highest remaining missing percentages. This residual missingness reflects prolonged station outages and is intentionally retained to preserve realistic data structure for mixed-frequency modelling.

Overall, the EDA confirms that the MBIE and NIWA datasets exhibit strong seasonality, realistic variability, and manageable data imperfections. The preprocessing strategy ensures complete quarterly electricity generation series while preserving the higher-frequency structure of daily weather data, providing a robust empirical foundation for the forecasting models evaluated in subsequent sections.

3.2 Baseline time-series forecasting results

This section reports the empirical performance of baseline univariate time-series forecasting models applied to quarterly net electricity generation in New Zealand. These

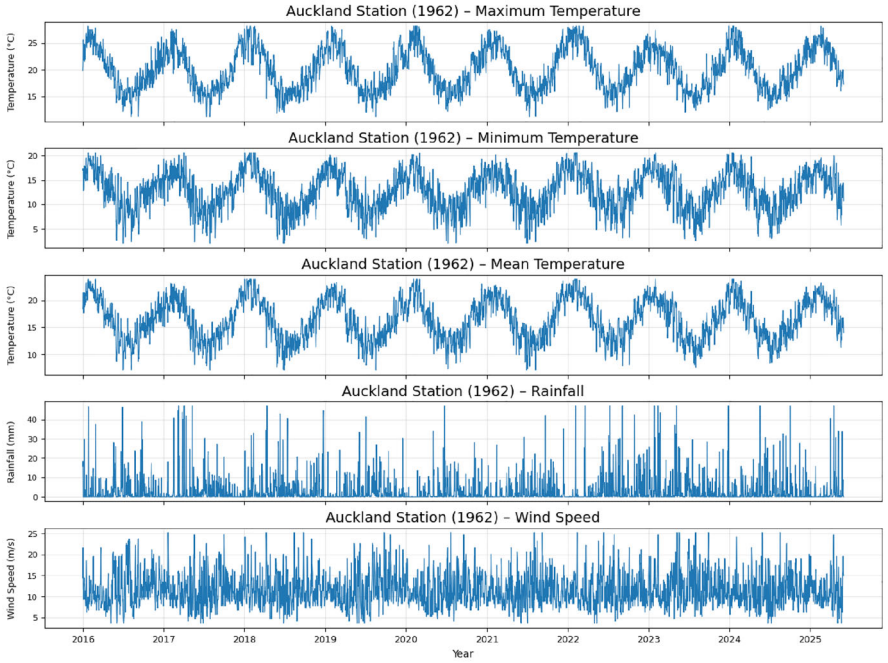


Fig. 5 Daily weather series for selected NIWA stations over the last ten years

Table 1 Top ten NIWA daily weather variables by remaining missing percentage after preprocessing

Variable	Missing (%)
11104_tmax	14.39
11104_tmin	14.39
11104_tmean	14.39
11104_rain	14.36
11104_wind	14.36
8567_wind	9.15
8567_tmean	9.12
8567_tmin	9.12
8567_tmax	9.12
7339_rain	6.35

baseline models rely solely on historical generation data and exclude weather or other exogenous information. The primary role is to provide transparent benchmark forecasts against which the benefits of mixed-frequency MIDAS models and machine learning extensions can be systematically evaluated.

Baseline time-series methods remain widely used in electricity forecasting due to their simplicity, interpretability, and strong performance in settings where demand exhibits stable seasonal patterns (Hyndman & Athanasopoulos, 2018; Shumway & Stoffer, 2017). In particular, quarterly electricity generation data are characterised

Table 2 Forecast accuracy of the baseline models over the last 8 quarters (rolling-origin evaluation)

Model	RMSE (GWh)	MAE (GWh)	MAPE (%)	MASE
Naïve	768.51	669.10	6.13	2.93
Seasonal Naïve ($m = 4$)	277.16	228.53	2.13	1.00
ARIMA (AIC-grid)	264.87	223.07	2.05	0.98
SARIMA (AIC-grid)	**201.33**	**173.63**	**1.58**	**0.76**
ETS (AIC-grid)	203.72	183.79	1.69	0.81

Note: Bold fonts indicate the best accuracy model

by pronounced seasonality and persistence, making classical linear models a natural starting point (Hahn et al., 2009; Paul et al., 2018).

Five baseline models are examined: the naïve, seasonal naïve, ARIMA, seasonal ARIMA (SARIMA), and exponential smoothing (ETS). These models represent increasing levels of structural complexity, ranging from simple persistence-based forecasts to models that explicitly capture trend and seasonal dynamics. All forecasts are generated using an expanding-window rolling-origin evaluation scheme consistent with real-time forecasting practice, ensuring that only information available at the time of forecast generation is used.

Forecast accuracy is assessed using standard error metrics commonly adopted in the forecasting literature, with comparative results presented later in this section and summarised across models. While baseline models are expected to capture broad seasonal and temporal patterns in electricity generation, they are inherently limited in their ability to incorporate higher-frequency weather information or non-linear relationships. These limitations motivate the transition to mixed-frequency and machine learning MIDAS models in subsequent sections.

Table 2 presents the forecast errors of the baseline models for the eight-quarter evaluation window. The SARIMA shows the best performance across accuracy measures with MASE below one, followed by the ETS. This indicates a substantial gain of SARIMA over the seasonal naïve and demonstrates that combining autoregressive dynamics with seasonal structure provides a strong univariate baseline. Nevertheless, remaining discrepancies suggest that important exogenous drivers, particularly higher-frequency weather variability, are not fully captured within a purely univariate framework, motivating the mixed-frequency models examined in the next section. For an illustrative purpose, Figs. 6, 7, 8, 9 and 10 show the forecasting performance of the five baseline models, which confirms the results in Table 2.

3.3 Classical MIDAS regression results

The classical MIDAS regression model is used to link the daily weather information to quarterly net electricity generation. Daily mean temperature, rainfall, and wind speed from NIWA stations are incorporated as higher-frequency predictors, while quarterly net generation from MBIE serves as the low-frequency target variable. Fore-

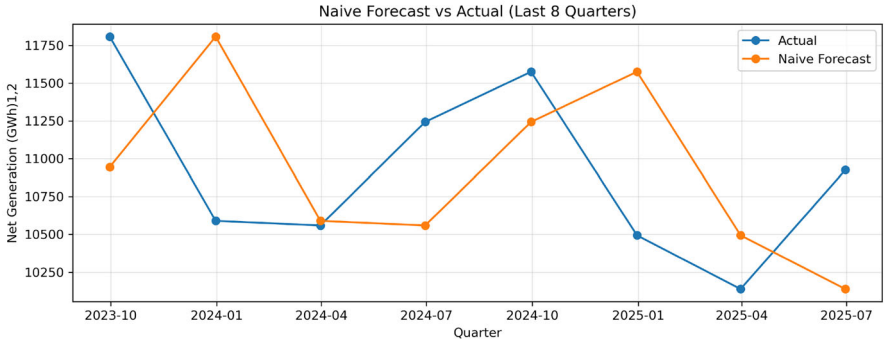


Fig. 6 Naïve one-step-ahead forecasts compared with actual quarterly net generation over the last eight quarters

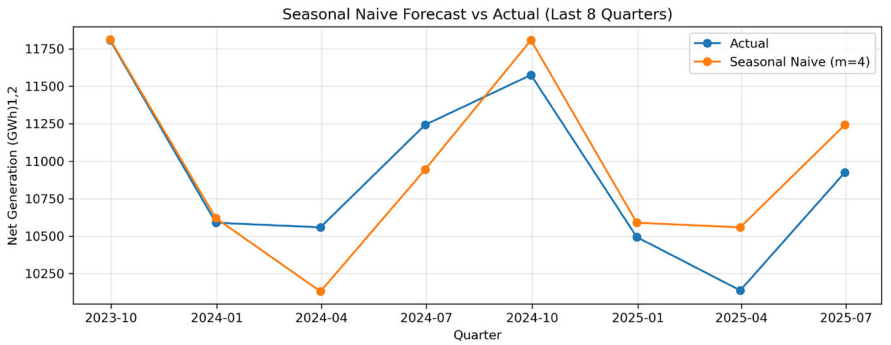


Fig. 7 Seasonal naïve one-step-ahead forecasts ($m = 4$) compared with actual quarterly net generation over the last eight quarters

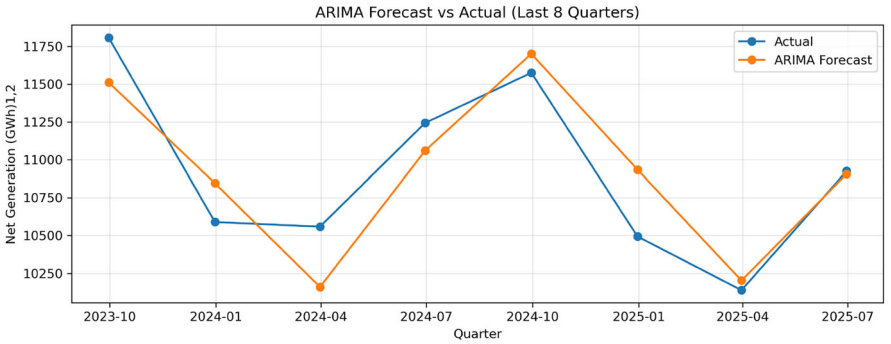


Fig. 8 One-step-ahead ARIMA forecasts compared with actual quarterly net electricity generation over the last eight quarters

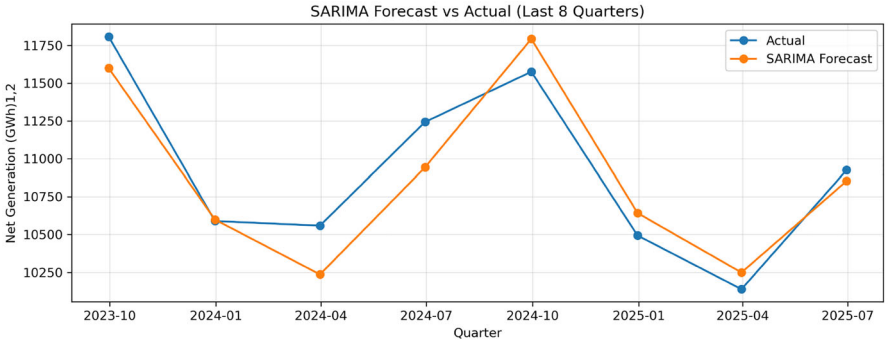


Fig. 9 One-step-ahead SARIMA forecasts compared with actual quarterly net electricity generation over the last eight quarters

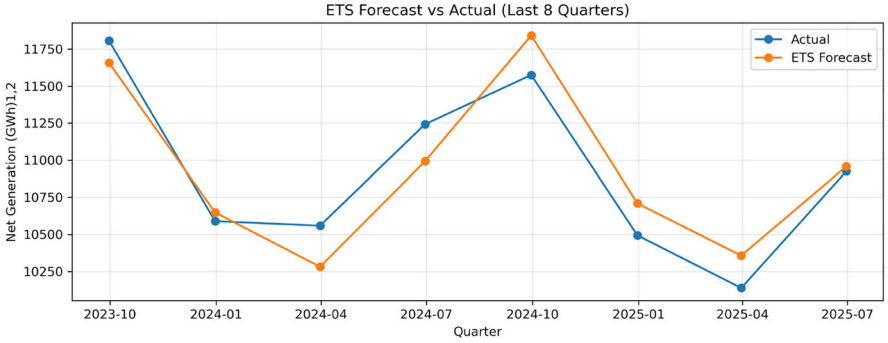


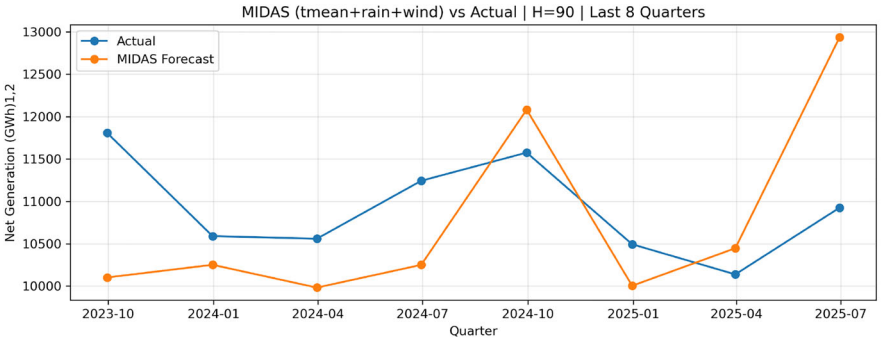
Fig. 10 One-step-ahead ETS forecasts compared with actual quarterly net electricity generation over the last eight quarters

cast accuracy is evaluated using an expanding-window rolling scheme over the final eight quarters of the sample.

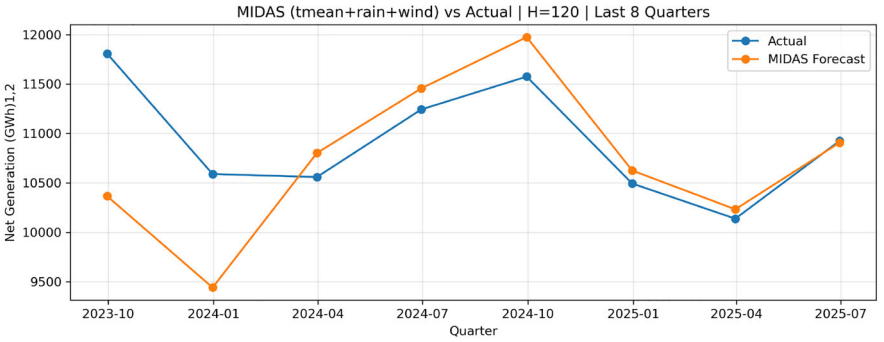
To examine the sensitivity of the MIDAS model to the aggregation window, three alternative lag lengths are considered: $H = 90$, $H = 120$, and $H = 150$ days prior to each quarterly observation. Figure 11 compares the resulting forecasts against observed quarterly electricity generation, while Table 3 reports the corresponding out-of-sample accuracy measures.

Across the three configurations, the MIDAS model with a 120-day aggregation window provides the most accurate and stable forecasts. This specification consistently achieves lower forecast errors than both shorter ($H = 90$) and longer ($H = 150$) windows, indicating that a moderate aggregation horizon best captures the cumulative influence of weather conditions on electricity generation. In contrast, the $H = 90$ model exhibits larger deviations in several quarters, suggesting that shorter windows may fail to capture persistent weather effects. The $H = 150$ specification shows signs of over-smoothing, leading to reduced responsiveness during periods of rapid change.

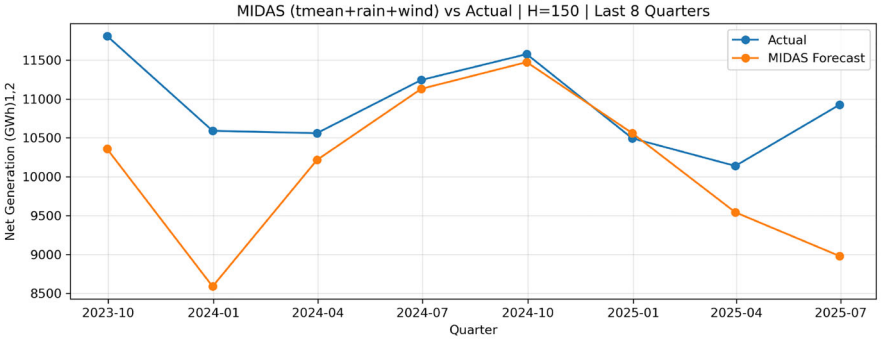
The performance differences reported in Table 3 confirm the visual evidence from Fig. 11. The $H = 120$ MIDAS model substantially reduces all error measures rela-



(a) $H = 90$ days



(b) $H = 120$ days



(c) $H = 150$ days

Fig. 11 Classical MIDAS forecasts versus actual quarterly net electricity generation for alternative aggregation horizons

Table 3 Forecasting accuracy of classical MIDAS models for alternative aggregation horizons

Model	Horizon (H)	RMSE (GWh)	MAE (GWh)	MAPE (%)	MASE
MIDAS (tmean + rain + wind)	90	1059.65	866.44	7.81	3.80
MIDAS (tmean + rain + wind)	120	679.02	461.26	4.13	2.02
MIDAS (tmean + rain + wind)	150	1140.20	827.85	7.59	3.63

tive to the alternative specifications, demonstrating improved generalisation across the evaluation period. However, despite incorporating higher-frequency weather information, the classical MIDAS models remain more volatile than the strongest univariate seasonal benchmarks reported earlier. This suggests that additional regularisation or nonlinear structure may be required to fully stabilise mixed-frequency forecasts in this setting.

These findings motivate the extension to advanced and machine-learning-based MIDAS models, which are examined in the subsequent section.

3.4 Advanced MIDAS results

While the classical MIDAS framework provides a useful baseline for incorporating higher-frequency weather information into quarterly electricity demand forecasts, its performance may be affected by multicollinearity, parameter instability, and limited flexibility when many higher-frequency predictors are included (Foroni et al., 2015; Ghysels et al., 2007). These issues are particularly relevant in the present study, where multiple station-level daily weather variables are aggregated into quarterly forecasts.

To address these limitations, this section extends the analysis to advanced MIDAS specifications that introduce regularisation and dynamic structure. Specifically, two advanced MIDAS variants are considered. The first is Ridge-MIDAS, which incorporates ℓ_2 regularisation into the MIDAS regression to stabilise coefficient estimation and mitigate the effects of multicollinearity among higher-frequency regressors (Hoerl & Kennard, 1970; Zou & Hastie, 2005). The second is AR-MIDAS, which augments the MIDAS regression with autoregressive terms of quarterly electricity generation, allowing persistent low-frequency dynamics to be captured alongside weather-driven effects (Andreou et al., 2013).

The forecasting performance of these advanced MIDAS models is evaluated using the same expanding-window rolling procedure and evaluation period as in the classical MIDAS analysis. This consistent evaluation framework enables direct comparison across baseline time-series models, classical MIDAS specifications, and advanced MIDAS extensions.

Table 4 summarises the out-of-sample forecasting performance of Ridge-MIDAS and AR-MIDAS over the final eight quarters. AR-MIDAS clearly outperforms Ridge-MIDAS in all four accuracy measures, with substantially lower RMSE, MAE, MAPE, and MASE values. This result shows that adding autoregressive structure to the MIDAS framework provides a major improvement in forecasting performance. It also suggests

Table 4 Ridge and AR-MIDAS forecasting performance over the last eight quarters ($H = 120$)

Model	RMSE (GWh)	MAE (GWh)	MAPE (%)	MASE
Ridge-MIDAS	679.03	461.26	4.13	2.02
AR-MIDAS	**214.42**	**157.15**	**1.42**	**0.69**

Note: Bold fonts indicate the best accuracy model

that persistence in quarterly electricity generation is an important source of predictive information in this setting.

The estimated autoregressive coefficient in the rolling AR-MIDAS forecasts remained stable across the eight folds, ranging from 0.702 to 0.730, with an average of approximately 0.712. This indicates that the recent history of quarterly electricity generation contains substantial predictive information beyond that captured by higher-frequency weather variables alone. From a statistical perspective, AR-MIDAS combines two useful components: autoregressive dependence in the low-frequency target and distributed lag effects from higher-frequency weather predictors. The strong performance of AR-MIDAS relative to Ridge-MIDAS and the ML-MIDAS models suggests that temporal persistence is an important source of predictive information in this quarterly forecasting framework.

To provide further insight into the implied lag weighting structure, Table 5 reports the estimated Beta shape parameters from each of the eight rolling AR-MIDAS forecast folds. These estimates help explain the memory effect of weather variables on the quarterly target and show how the implied weighting pattern changes only slightly across the rolling forecast folds. In particular, the results suggest that rainfall and especially wind place greater emphasis on more recent observations, while temperature effects are somewhat more distributed over recent to intermediate lags. This indicates that the influence of weather variables decays over time, but the rate of decay differs across weather types.

The results from Classical MIDAS and Ridge-MIDAS are almost identical, which suggests that in this case the ridge penalty does not provide an additional forecasting gain beyond the baseline MIDAS structure. By contrast, the much stronger performance of AR-MIDAS shows that accounting for persistence in the quarterly target is more valuable than regularisation alone. Figures 12 and 13 are presented for comparison and completeness.

3.5 Machine learning MIDAS results

While econometric MIDAS specifications provide an interpretable and structured framework for incorporating mixed-frequency weather information, their linear structure may limit their ability to capture complex nonlinear relationships between weather drivers and quarterly electricity generation. To address this limitation, we extend the analysis to Machine Learning MIDAS (ML-MIDAS) models, which combine MIDAS-style mixed-frequency feature construction with flexible nonlinear learning algorithms that can model interactions and nonlinear effects more effectively (Bon-

Table 5 Estimated Beta shape parameters from the rolling AR-MIDAS forecasts

Fold	Test quarter	ϕ	a_{temp}	b_{temp}	a_{rain}	b_{rain}	a_{wind}	b_{wind}
1	2023Q3	0.730	3.019	2.468	1.568	2.787	1.490	3.682
2	2023Q4	0.714	2.383	1.793	1.380	2.430	1.489	3.563
3	2024Q1	0.709	2.295	1.712	1.349	2.403	1.474	3.487
4	2024Q2	0.707	2.331	1.731	1.369	2.424	1.474	3.459
5	2024Q3	0.708	2.373	1.766	1.381	2.445	1.481	3.493
6	2024Q4	0.712	2.412	1.811	1.401	2.521	1.479	3.468
7	2025Q1	0.711	2.411	1.806	1.401	2.515	1.478	3.457
8	2025Q2	0.702	2.261	1.689	1.345	2.461	1.484	3.347

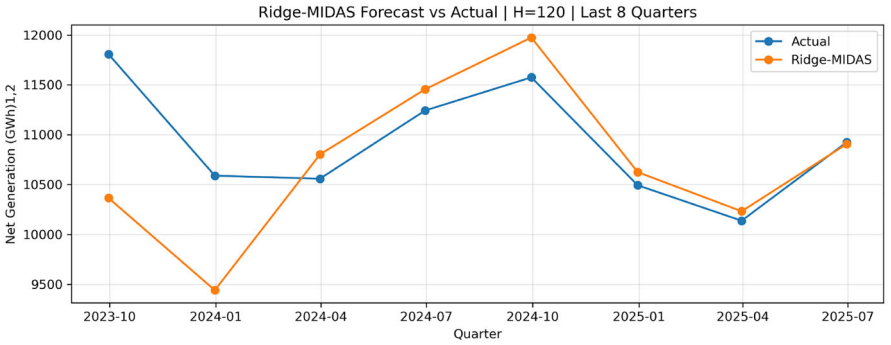


Fig. 12 Ridge-MIDAS forecast versus actual quarterly net generation over the last eight quarters ($H = 120$)

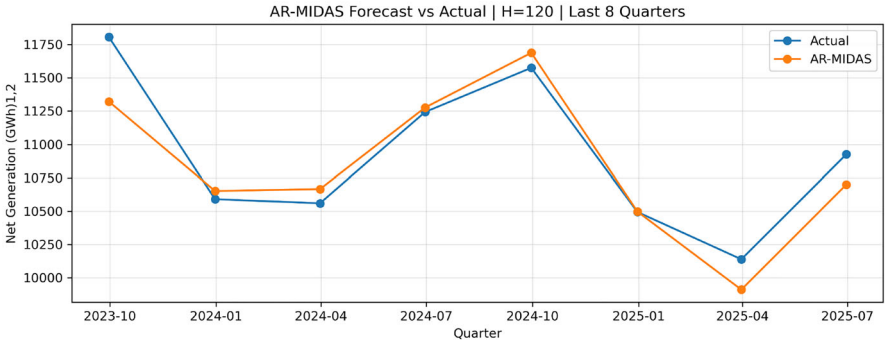


Fig. 13 AR-MIDAS forecast versus actual quarterly net electricity generation over the last eight quarters ($H = 120$)

tempi et al., 2013; Ghysels et al., 2007). This approach is consistent with the broader forecasting literature that highlights the benefits of machine learning methods for complex time series prediction problems, including energy applications (Breiman, 2001; Friedman, 2001; Ibrahim et al., 2022). The forecasting performance of the ML-MIDAS models is evaluated using the same expanding-window rolling proce-

Table 6 Forecast accuracy for the ML-MIDAS models over the last eight quarters ($H = 120$)

Model	RMSE (GWh)	MAE (GWh)	MAPE (%)	MASE
LASSO-MIDAS	592.59	479.82	4.46	2.10
ElasticNet-MIDAS	**581.68**	**462.14**	**4.29**	**2.03**
RF-MIDAS	779.89	696.24	6.29	3.05
GBR-MIDAS	748.58	644.71	5.80	2.83
XGB-MIDAS	822.14	697.09	6.26	3.06

Note: Bold fonts indicate the best accuracy model

dure and the same last-eight-quarters test period as the econometric MIDAS models, ensuring a fair and directly comparable assessment across model classes that include LASSO-MIDAS, ElasticNet-MIDAS, Random Forest-MIDAS (RF-MIDAS), Gradient Boosting-MIDAS (GBR-MIDAS), and XGBoost-MIDAS.

Table 6 summarises the out-of-sample forecasting performance of all ML-MIDAS models used in this study. The results show that the ElasticNet-MIDAS is the best model in all measures, followed by the LASSO-MIDAS. However, these ML-MIDAS models are still outperformed by the AR-MIDAS and most of the baseline models, except the naïve model. Note that the graphical results from the ML-MIDAS models are omitted for conciseness.

One possible reason why the ML-MIDAS models did not outperform AR-MIDAS is that the higher-frequency weather information was compressed using fixed Beta weights before the machine learning stage. This choice was made for tractability, as it keeps the mixed-frequency feature construction simple and computationally feasible in a rolling forecasting setting. However, it may also reduce flexibility, because the machine learning models do not learn the lag-weighting structure directly from the data. Instead, they operate on pre-aggregated MIDAS features whose temporal structure has already been constrained by the fixed weighting scheme. In contrast, AR-MIDAS combines an explicit autoregressive component with optimised mixed-frequency lag effects, which may be better suited to the present quarterly forecasting problem. This suggests that, in this study, the main forecasting gain comes from capturing persistence and stable lagged weather effects, while the use of fixed Beta weights may have limited the ability of the ML-MIDAS models to fully exploit more flexible nonlinear relationships.

Overall, the comparison indicates that modelling temporal persistence explicitly within the MIDAS framework is important for improved medium-term electricity generation forecasts. AR-MIDAS performs strongly and clearly outperforms the other MIDAS variants, while SARIMA remains highly competitive and achieves the lowest RMSE in the main eight-quarter evaluation period. The ML-MIDAS approaches do not consistently outperform the strongest advanced MIDAS or seasonal benchmark models in this setting.

Table 7 Simulation design and parameter settings

Component	Setting
Number of replications	100
Number of quarterly observations	120
Test horizon	Last 8 quarters
Seasonal period	4
MIDAS horizon (H)	120
Intercept (α)	50.0
Autoregressive coefficient (ϕ)	0.60
Temperature coefficient (β_{temp})	1.20
Rainfall coefficient (β_{rain})	−0.80
Wind coefficient (β_{wind})	0.60
Error standard deviation (σ_y)	1.0
Daily temperature persistence (ρ_{temp})	0.85
Daily rainfall persistence (ρ_{rain})	0.65
Daily wind persistence (ρ_{wind})	0.75
Quarterly seasonal effects	Q1: −1.0, Q2: 0.5, Q3: 1.2, Q4: 0.3
Models compared	SARIMA, Classical MIDAS, AR-MIDAS, ElasticNet-MIDAS

3.6 Simulation study

To support the empirical findings, a simulation study was conducted in a controlled mixed-frequency setting. The purpose of this exercise was to examine whether the estimation and forecasting procedures used in this paper behave sensibly when the data contain persistence, seasonality, and higher-frequency predictor effects. This simulation does not replace the real-data analysis. Instead, it provides additional evidence under a known data-generating process in a setting consistent with the mixed-frequency MIDAS framework considered in this study (Andreou et al., 2013; Ghysels et al., 2007).

In each replication, three daily predictor series were generated to represent higher-frequency weather variables, namely temperature, rainfall, and wind. A quarterly target series was then generated using an autoregressive term, a seasonal component, MIDAS-style weighted higher-frequency predictors, and a random error term. The simulation design follows the data-generating process

$$y_t = \alpha + \phi y_{t-1} + \text{seasonal effect} + \beta_{\text{temp}} z_{\text{temp},t} + \beta_{\text{rain}} z_{\text{rain},t} + \beta_{\text{wind}} z_{\text{wind},t} + \varepsilon_t,$$

where $z_{\text{temp},t}$, $z_{\text{rain},t}$, and $z_{\text{wind},t}$ denote MIDAS-style weighted aggregates of the previous $H = 120$ daily observations. The simulation design and parameter settings are summarised in Table 7.

The simulation was repeated over 100 replications. To assess whether the underlying parameter values can be recovered reasonably well, Table 8 reports the true parameter values together with the mean estimated values, standard deviations, and

Table 8 Parameter recovery results for the data-generating regression across 100 simulation replications

Parameter	True value	Mean estimated	SD estimated	MAD estimated
α	50.0	49.053	2.181	1.292
ϕ	0.6	0.599	0.017	0.011
β_{temp}	1.2	1.216	0.143	0.093
β_{rain}	-0.8	-0.791	0.309	0.231
β_{wind}	0.6	0.565	0.216	0.147
Season Q2 effect	1.5	1.498	0.308	0.204
Season Q3 effect	2.2	2.213	0.352	0.202
Season Q4 effect	1.3	1.322	0.412	0.258

Note: Seasonal effects are reported relative to Q1, which is the base quarter in the estimation

Table 9 Simulation study forecast accuracy based on 100 replications

	RMSE		MAE		MAPE (%)		MASE	
Model	Mean	SD	Mean	SD	Mean	SD	Mean	SD
AR-MIDAS ($H = 120$)	**1.250**	0.287	**1.056**	0.258	**0.842**	0.207	**0.416**	0.106
SARIMA	1.432	0.347	1.183	0.298	0.943	0.239	0.466	0.126
ElasticNet-MIDAS ($H = 120$)	2.166	0.625	1.834	0.554	1.452	0.425	0.725	0.238
Classical MIDAS ($H = 120$)	2.185	0.618	1.849	0.553	1.465	0.425	0.731	0.238

Note: Bold fonts indicate the best mean forecast accuracy

median absolute deviations across replications. The results show that the main parameters are recovered accurately on average. In particular, the autoregressive coefficient is estimated almost exactly, with a mean estimate of 0.599 compared with the true value of 0.600. The temperature coefficient is also recovered very closely, with a mean estimate of 1.216 for a true value of 1.200. The rainfall coefficient is estimated at -0.791, which is very close to the true value of -0.800, while the wind coefficient is slightly underestimated on average at 0.565 but still remains reasonably close to the true value of 0.600. The seasonal dummy effects are likewise recovered well, with mean estimated values of 1.498, 2.213, and 1.322 for the Q2, Q3, and Q4 effects, compared with their true values of 1.5, 2.2, and 1.3, respectively. Overall, these results suggest that the estimation procedure performs well in repeated samples and is able to recover the main persistence, seasonal, and mixed-frequency weather effects embedded in the data-generating process.

Four forecasting models were then compared in the simulation study: SARIMA, Classical MIDAS ($H = 120$), AR-MIDAS ($H = 120$), and ElasticNet-MIDAS ($H = 120$). The evaluation followed the same expanding-window rolling one-step-ahead forecasting design used in the empirical analysis, with the last eight quarters treated as the test period. Forecast accuracy was summarised using RMSE, MAE, MAPE, and MASE (Hyndman & Koehler, 2006). The results are reported in Table 9. AR-MIDAS achieved the best overall forecasting performance across all four accuracy measures, with the lowest mean RMSE (1.250), mean MAE (1.056), mean MAPE (0.842), and

mean MASE (0.416). SARIMA was the second-best model, while ElasticNet-MIDAS and Classical MIDAS produced noticeably larger forecast errors and very similar results to each other.

The results in Table 9 are important because they help explain why AR-MIDAS performs well in this controlled setting. The simulated data-generating process includes three important features simultaneously: autoregressive persistence in the quarterly target, seasonal variation across quarters, and mixed-frequency effects coming from the higher-frequency weather predictors. AR-MIDAS is the only model among the competing specifications that directly combines autoregressive structure with MIDAS-type mixed-frequency information. Because of this, it is better aligned with the true underlying structure of the simulated data. SARIMA can capture persistence and seasonality, but it does not use the higher-frequency weather information explicitly. Classical MIDAS and ElasticNet-MIDAS use the mixed-frequency weather predictors, but they do not include the autoregressive term in the same direct way as AR-MIDAS. This explains why AR-MIDAS produces the lowest forecast errors on average across the 100 replications.

Figure 14 presents simulated actual values together with one-step-ahead forecasts from AR-MIDAS and SARIMA for a representative simulation replication over the final eight quarters. The figure shows that both models are able to follow the broad movement of the simulated series, but AR-MIDAS generally remains closer to the realised values across most of the evaluation period. In particular, its forecasts track the direction and magnitude of the quarterly changes more consistently, while SARIMA shows larger deviations in several quarters. This visual evidence is consistent with the average forecasting results reported in Table 9, where AR-MIDAS outperforms SARIMA across all four mean accuracy measures. Taken together, the parameter recovery results, the forecasting summary, and the representative simulation plot provide further support for the usefulness of AR-MIDAS when the underlying data-generating process contains persistence, seasonality, and mixed-frequency weather effects.

3.7 Sensitivity analysis across evaluation periods

To examine whether the main forecasting conclusions depend on the selected evaluation period, a sensitivity analysis was conducted using alternative rolling test windows of 6, 8, 10, and 12 quarters. Forecast accuracy was assessed using the same error measures adopted throughout the study, namely RMSE, MAE, MAPE, and MASE (Hyndman & Koehler, 2006). Table 10 reports the corresponding results for SARIMA, Ridge-MIDAS, AR-MIDAS, and ElasticNet-MIDAS across these alternative evaluation windows.

The results show that the relative ranking of the leading models varies across evaluation periods. AR-MIDAS performs best over the shorter 6-quarter window and remains competitive over the 8-quarter period, where it achieves the lowest MAE, MAPE, and MASE, while SARIMA achieves the lowest RMSE. When the evaluation window is extended to 10 and 12 quarters, SARIMA becomes the strongest model across all four accuracy measures. These findings suggest that model conclusions are somewhat

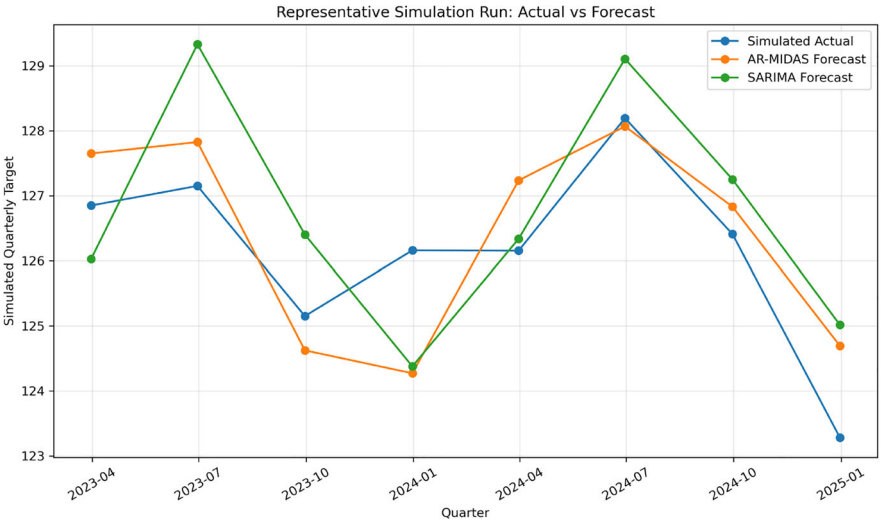


Fig. 14 Simulated actual values and one-step-ahead forecasts from AR-MIDAS and SARIMA for a representative simulation replication over the final eight quarters

sensitive to the chosen evaluation period. This also indicates that forecast rankings may be affected by period-specific conditions, including unusual weather patterns or broader structural changes in the electricity system, so the main results should not be interpreted as fully invariant across all sample periods.

To assess whether the observed differences in forecast accuracy are statistically significant, the Diebold-Mariano test for equal predictive accuracy was applied using squared forecast errors (Diebold & Mariano, 1995). Table 11 reports the corresponding test results. Although AR-MIDAS and SARIMA alternate as the lower-loss model across different evaluation windows, none of the pairwise differences are statistically significant at the 5% level. Likewise, the differences between AR-MIDAS and the other MIDAS variants are not statistically significant. Overall, these results suggest that while forecast rankings vary across evaluation periods, the observed differences in predictive accuracy should be interpreted with caution.

4 Discussion

This section provides an interpretative discussion of the findings and places them within the broader context of electricity demand forecasting and energy system planning in New Zealand. In particular, the discussion examines why mixed-frequency models that explicitly account for temporal persistence and higher-frequency weather information can improve forecasting performance relative to purely statistical baselines in some evaluation settings. The implications of incorporating regularisation, autoregressive structure, and machine learning techniques into the MIDAS frame-

Table 10 Sensitivity analysis of forecast accuracy across alternative evaluation windows

Test horizon	Model	RMSE	MAE	MAPE (%)	MASE
6	AR-MIDAS	**146.21**	**118.33**	**1.10**	**0.53**
6	SARIMA	216.34	195.17	1.79	0.87
6	Ridge-MIDAS	220.09	183.29	1.67	0.81
6	ElasticNet-MIDAS	382.98	322.32	3.02	1.43
8	SARIMA	**201.33**	173.63	1.58	0.76
8	AR-MIDAS	214.42	**157.15**	**1.42**	**0.69**
8	ElasticNet-MIDAS	581.68	462.14	4.29	2.03
8	Ridge-MIDAS	679.02	461.26	4.13	2.02
10	SARIMA	**191.66**	**167.74**	**1.53**	**0.73**
10	AR-MIDAS	423.89	257.17	2.34	1.12
10	ElasticNet-MIDAS	543.32	439.40	4.09	1.91
10	Ridge-MIDAS	698.43	520.41	4.73	2.27
12	SARIMA	**192.15**	**165.26**	**1.50**	**0.72**
12	AR-MIDAS	459.73	314.50	2.84	1.37
12	Ridge-MIDAS	666.26	509.12	4.63	2.21
12	ElasticNet-MIDAS	694.71	555.61	5.08	2.41

Note: Bold fonts indicate the best accuracy model within each evaluation window

Table 11 Diebold-Mariano test results across alternative evaluation windows

Horizon	Comparison	DM stat	p-value	Better model	Sig. (5%)
6	AR-MIDAS vs SARIMA	−0.947	0.344	AR-MIDAS	No
6	AR-MIDAS vs Ridge-MIDAS	−0.837	0.402	AR-MIDAS	No
6	AR-MIDAS vs ElasticNet-MIDAS	−1.575	0.115	AR-MIDAS	No
8	AR-MIDAS vs SARIMA	0.157	0.875	SARIMA	No
8	AR-MIDAS vs Ridge-MIDAS	−1.495	0.135	AR-MIDAS	No
8	AR-MIDAS vs ElasticNet-MIDAS	−1.446	0.148	AR-MIDAS	No
10	AR-MIDAS vs SARIMA	0.967	0.333	SARIMA	No
10	AR-MIDAS vs Ridge-MIDAS	−1.280	0.200	AR-MIDAS	No
10	AR-MIDAS vs ElasticNet-MIDAS	−0.501	0.616	AR-MIDAS	No
12	AR-MIDAS vs SARIMA	1.415	0.157	SARIMA	No
12	AR-MIDAS vs Ridge-MIDAS	−1.128	0.259	AR-MIDAS	No
12	AR-MIDAS vs ElasticNet-MIDAS	−1.149	0.250	AR-MIDAS	No

work are evaluated, with attention given to model robustness, interpretability, and practical deployment considerations.

The discussion starts with the implications of the empirical findings for electricity demand forecasting in New Zealand, particularly in the context of increasing renewable penetration and weather sensitivity (Ministry of Business, 2019, 2023). Then, we compare the observed model behaviour with existing literature on mixed-frequency

forecasting and machine learning-based load prediction (Clements & Galvão, 2008; Ghysels et al., 2007; Hong & Fan, 2016). After that, we outline key limitations of the present study and highlight directions for future research.

4.1 Implications for electricity demand forecasting in New Zealand

The findings of this study have important implications for electricity demand forecasting in New Zealand, where demand patterns are increasingly shaped by weather variability and the ongoing transition towards a highly renewable electricity system. As highlighted in national energy policy documents, accurate demand forecasts are essential for system planning, operational reliability, and investment decision-making in a power system dominated by hydro, wind, and other weather-dependent generation sources (Ministry of Business, 2019, 2023).

The results indicate that incorporating higher-frequency weather information can improve quarterly electricity demand forecasts relative to traditional univariate time-series approaches, although the relative gains depend on the selected model and evaluation period. This result aligns with existing evidence that temperature, rainfall, and wind conditions exert a significant influence on electricity demand through both direct consumption effects and indirect supply-side dynamics (Mirasgedis et al., 2006; Pardo et al., 2002; Taylor & Buizza, 2003). In the New Zealand context, where hydro inflows and heating demand are closely linked to climatic conditions, such mixed-frequency integration is particularly relevant.

Among the models evaluated, the strong performance of AR-MIDAS highlights the importance of jointly modelling weather effects and temporal persistence in electricity demand. The inclusion of an autoregressive component allows the model to capture structural inertia and long-term consumption trends that are not fully explained by weather variables alone. However, the sensitivity analysis also shows that the relative ranking of AR-MIDAS and SARIMA depends on the evaluation period, with AR-MIDAS performing best over the shorter recent window and SARIMA becoming stronger when the test period is extended. This finding is consistent with prior studies emphasising the role of dynamic dependence in medium-term electricity forecasting (Hahn et al., 2009; Shumway & Stoffer, 2017).

The results further suggest that regularisation-based approaches, such as Ridge-MIDAS and ML-MIDAS variants, offer practical advantages in high-dimensional forecasting environments. By stabilising coefficient estimates and mitigating multicollinearity among weather predictors, these models enhance forecast robustness while retaining flexibility in capturing nonlinear relationships. Such characteristics are particularly valuable for operational forecasting systems that rely on large-scale meteorological datasets (Hoerl & Kennard, 1970; Hong & Fan, 2016).

From a policy and operational perspective, these findings support the adoption of advanced mixed-frequency forecasting frameworks within New Zealand's electricity sector. Improved demand forecasts can assist grid operators and policymakers in managing supply adequacy, integrating renewable resources, and assessing climate-related risks. As the energy transition accelerates, forecasting approaches that explicitly account for weather-driven uncertainty are likely to play an increasingly central role

in ensuring system resilience and efficiency (Ministry of Business, 2024; National Institute of Water and Atmospheric Research, 2024).

4.2 Comparison with existing literature

The results of this study are broadly consistent with the existing literature on electricity demand forecasting, while also extending prior findings through the explicit use of mixed-frequency and advanced MIDAS frameworks in a New Zealand context. Traditional studies have long established the importance of weather variables, particularly temperature, in explaining electricity demand dynamics (Mirasgedis et al., 2006; Pardo et al., 2002). The improved performance of weather-augmented models observed in this study reinforces these conclusions and highlights the continued relevance of meteorological drivers in medium-term forecasting.

Compared to conventional time-series approaches such as ARIMA, SARIMA, and exponential smoothing, the mixed-frequency models examined here demonstrate clear advantages. Earlier work has shown that pure time-series models, while effective for short-term forecasting, often struggle to capture structural changes and exogenous influences at longer horizons (Gardner, 2006; Ho & Xie, 1998). The findings of this study support this view, as models that rely solely on historical demand patterns exhibit inferior predictive accuracy when compared with MIDAS-based specifications that incorporate daily weather information.

The strong performance of classical MIDAS models aligns with prior macroeconomic and energy forecasting studies that emphasise the value of mixed-frequency data integration (Clements & Galvão, 2008; Ghysels et al., 2007). By allowing higher-frequency weather variables to inform lower-frequency demand outcomes, MIDAS provides a flexible and theoretically grounded framework for addressing frequency mismatches commonly encountered in energy datasets. This study extends these insights by demonstrating the applicability of MIDAS to electricity demand forecasting in a small, renewable-dominated power system.

Beyond classical MIDAS, the results for AR-MIDAS suggest that incorporating autoregressive dynamics alongside mixed-frequency weather effects can substantially improve forecast accuracy, particularly over shorter recent evaluation windows. Previous research has shown that electricity demand exhibits strong temporal persistence due to behavioural, technological, and structural factors (Hahn et al., 2009; Shumway & Stoffer, 2017). Incorporating lagged demand alongside weather-driven MIDAS terms enables the model to capture both short-term climatic effects and longer-term consumption inertia, leading to improved forecast performance.

The mixed results obtained from machine learning-based MIDAS models are also in line with earlier findings in the energy forecasting literature. While machine learning methods such as random forests, gradient boosting, and regularised regression have shown promise in capturing nonlinear relationships (Breiman, 2001; Friedman, 2001; Hong & Fan, 2016), their performance can be sensitive to sample size, tuning choices, and forecast horizon. In the present study, linear regularised models such as LASSO-MIDAS and ElasticNet-MIDAS perform competitively, whereas more com-

plex tree-based models offer limited gains, reflecting observations reported in previous comparative studies (Ibrahim et al., 2022; Lahouar & Ben Hadj Slama, 2015).

Overall, this study contributes to the existing body of literature by providing empirical evidence that advanced MIDAS models, particularly those incorporating autoregressive structure and regularisation, offer a robust and interpretable alternative to both traditional time-series models and purely data-driven machine learning approaches. These findings support recent calls for hybrid modelling strategies that balance statistical rigour, interpretability, and predictive accuracy in electricity demand forecasting applications (Bontempi et al., 2013; Weron, 2014).

4.3 Limitations and future research directions

While this study demonstrates the effectiveness of mixed-frequency and advanced MIDAS models for electricity demand forecasting in New Zealand, several limitations should be acknowledged, which also point to directions for future research.

First, the evaluation period used for model comparison is limited to the most recent eight quarters. Although an expanding-window framework was adopted to ensure robustness and avoid information leakage, the relatively short out-of-sample horizon may not fully capture performance under extreme events or structural breaks, such as prolonged droughts, major policy interventions, or sudden changes in generation mix. Future work could extend the evaluation window or conduct stress-testing under simulated extreme weather and demand scenarios.

Second, the set of explanatory variables in the MIDAS framework is restricted to daily weather indicators, namely temperature, rainfall, and wind. While these variables are well-established drivers of electricity demand, other factors such as electricity prices, industrial activity, electrification trends, and policy-driven behavioural changes were not explicitly modelled. Incorporating additional higher-frequency economic or system-level variables could further enhance forecast accuracy and provide deeper insights into demand dynamics.

Third, the machine learning-based MIDAS models rely on fixed beta-weight aggregation for computational tractability. Although this approach improves stability and runtime efficiency, it constrains the flexibility of the lag structure and may limit the ability of ML models to fully exploit temporal patterns in higher-frequency data. Future research could explore adaptive or learned aggregation schemes, including neural network-based MIDAS representations, while carefully managing overfitting risks.

Fourth, model interpretability remains an important consideration, particularly for operational and policy-oriented applications. While advanced models such as AR-MIDAS offer improved accuracy, their increased complexity may reduce transparency relative to simpler baseline approaches. Future studies could focus on developing explainability tools for MIDAS and ML-MIDAS models, such as feature attribution methods or scenario-based sensitivity analysis, to support decision-making by system operators and policymakers.

Finally, this study focuses on quarterly electricity demand forecasting. Extending the proposed framework to other temporal resolutions, such as monthly or semi-annual

horizons, or adapting it for forecasting peak demand and system adequacy metrics, represents a promising avenue for future research. Such extensions would further enhance the practical relevance of mixed-frequency modelling approaches in energy system planning and renewable integration contexts.

Overall, addressing these limitations would contribute to a more comprehensive and flexible forecasting framework, strengthening the role of mixed-frequency and hybrid models in supporting the ongoing energy transition in New Zealand.

5 Conclusion

This study develops and evaluates a comparative mixed-frequency forecasting framework for quarterly electricity demand in New Zealand, integrating daily weather information with advanced econometric and machine learning models. By combining baseline time-series approaches, classical MIDAS regression, advanced MIDAS variants, and machine learning-based MIDAS models, the analysis provides a systematic comparison of forecasting performance across a wide range of modelling strategies.

The results demonstrate that traditional univariate benchmarks such as SARIMA and ETS offer reasonable baseline performance but are limited in their ability to fully exploit higher-frequency weather information. Classical MIDAS regression improves forecast accuracy by explicitly incorporating daily weather drivers; however, its performance is sensitive to lag selection and parameter stability. Among the classical MIDAS specifications, the horizon length of 120 days provides the most consistent results, highlighting the importance of capturing medium-term weather effects.

Advanced MIDAS models deliver substantial improvements over classical MIDAS specifications. In particular, AR-MIDAS provides strong forecasting performance by combining autoregressive dynamics with mixed-frequency weather information, and it clearly outperforms the other MIDAS variants considered in this study. However, the comparison with seasonal benchmark models is more nuanced. In the main eight-quarter evaluation period, SARIMA achieves the lowest RMSE, while AR-MIDAS performs better on MAE, MAPE, and MASE. Additional sensitivity analysis further shows that AR-MIDAS performs best over the shorter recent evaluation window, whereas SARIMA becomes stronger when the evaluation period is extended. These findings suggest that the relative ranking of the leading models depends on the chosen evaluation horizon.

Machine learning-based MIDAS models further demonstrate the potential of flexible, data-driven approaches in mixed-frequency settings. Regularised linear models, especially ElasticNet-MIDAS and LASSO-MIDAS, perform competitively by effectively handling high-dimensional weather inputs while maintaining generalisation. In contrast, tree-based models such as Random Forest, Gradient Boosting, and XGBoost exhibit weaker performance in this application, suggesting that their strengths may be better suited to larger sample sizes or shorter-term forecasting horizons.

Overall, the findings highlight the value of integrating higher-frequency climate information with both econometric structure and modern regularisation techniques. The proposed framework offers a robust and scalable approach for medium-term electricity demand forecasting, with direct relevance for system planning, renewable

integration, and policy analysis in New Zealand's transitioning energy system. The results also underscore the broader applicability of mixed-frequency modelling techniques for energy forecasting problems where data arrive at heterogeneous temporal resolutions.

Acknowledgements The authors acknowledge the industry feedback and contextual guidance provided through collaboration with Mercury Energy New Zealand Limited. This collaboration supported the practical framing of the research problem. Only publicly available data from the Ministry of Business, Innovation and Employment (MBIE) and the National Institute of Water and Atmospheric Research (NIWA) were used, and no proprietary or confidential information was accessed or included in this study.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions This project received the financial support from the 2025–2026 DCT Summer Research Scholarships scheme.

Data Availability All data used in this study are publicly available from MBIE and NIWA.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Amjady, N. (2001). Short-term hourly load forecasting using time-series modeling with peak load estimation capability. *IEEE Transactions on Power Systems, 16*(4), 798–805. https://doi.org/10.1109/59.962429

Andreou, E., Ghysels, E., & Kourtellos, A. (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics, 31*(2), 240–251. https://doi.org/10.1080/07350015.2013.767199

Bessec, M., & Fouquau, J. (2008). The non-linear link between electricity consumption and temperature in europe. *Energy Economics, 30*(5), 2705–2721. https://doi.org/10.1016/j.eneco.2008.02.003

Bontempi, G., Ben Taieb, S., & Le Borgne, Y.-A. (2013). Machine learning strategies for time series forecasting. *Business Intelligence* (pp. 62–77). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-36318-4_3

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B, 26*(2), 211–252. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32. https://doi.org/10.1023/A:1010933404324

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Clements, M. P., & Galvão, A. B. (2008). Macroeconomic forecasting with mixed-frequency data: Forecasting us output growth. *Journal of Business & Economic Statistics, 26*(4), 546–554. https://doi.org/10.1198/073500108000000015

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 13*(3), 253–263. https://doi.org/10.1080/07350015.1995.10524599

Foroni, C., Marcellino, M., & Schumacher, C. (2015). Unrestricted mixed data sampling (midas): Midas regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society, Series A, 178*(1), 57–82. https://doi.org/10.1111/rssa.12043

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Gardner, E. S. (2006). Exponential smoothing: The state of the art - part ii. *International Journal of Forecasting, 22*(4), 637–666. https://doi.org/10.1016/j.ijforecast.2006.03.005

Ghelasi, P., & Ziel, F. (2025). From day-ahead to mid and long-term horizons with econometric electricity price forecasting models. *Renewable and Sustainable Energy Reviews, 217*, Article 115684. https://doi.org/10.1016/j.rser.2025.115684

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). Predicting volatility: Getting the most out of return data sampled at different frequencies. *Journal of Econometrics, 131*(1–2), 59–95. https://doi.org/10.1016/j.jeconom.2005.01.004

Ghysels, E., Sinko, A., & Valkanov, R. (2007). Midas regressions: Further results and new directions. *Econometric Reviews, 26*(1), 53–90. https://doi.org/10.1080/07474930600972467

Hahn, H., Meyer-Nieberg, S., & Pickl, S. (2009). Electric load forecasting methods: Tools for decision making. *European Journal of Operational Research, 199*(3), 902–907. https://doi.org/10.1016/j.ejor.2009.01.062

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488634

Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting, 32*(3), 914–938. https://doi.org/10.1016/j.ijforecast.2015.11.011

Ho, S. L., & Xie, M. (1998). The use of arima models for reliability forecasting and analysis. *Computers & Industrial Engineering, 35*(1–2), 213–216. https://doi.org/10.1016/S0360-8352(98)00066-7

Hyndman, R.J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice, 2nd edn. OTexts, Melbourne, Australia. https://otexts.com/fpp2/

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*(4), 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Ibrahim, B., Rabelo, L., Gutierrez-Franco, E., & Clavijo-Buritica, N. (2022). Machine learning for short-term load forecasting in smart grids. *Energies, 15*(21), 8079. https://doi.org/10.3390/en15218079

Kapoor, G., Wichitaksorn, N., Li, M., & Zhang, W. (2025). Forecasting half-hourly electricity prices using a mixed-frequency structural VAR framework. *Econometrics, 13*(1), 2. https://doi.org/10.3390/econometrics13010002

Lahouar, A., & Ben Hadj Slama, J. (2015). Day-ahead load forecast using random forest and expert input selection. *Energy Conversion and Management, 103*, 1040–1051. https://doi.org/10.1016/j.enconman.2015.07.041

Little, R. J. A., & Rubin, D. B. (2019). Statistical Analysis with Missing Data. Wiley. https://doi.org/10.1002/9781119482260 3rd edn.

Magalhães, B. (2024). Short-term load forecasting based on optimized random forest and feature selection. *Energies, 17*(8), 1926. https://doi.org/10.3390/en17081926

Ministry of Business. (2019). Innovation and Employment: Energy in New Zealand 2019. https://www.mbie.govt.nz/assets/energy-in-new-zealand-2019.pdf

Ministry of Business. (2023). Innovation and Employment: Advancing New Zealand's energy transition consultation document. https://www.mbie.govt.nz/building-and-energy/energy-and-natural-resources/energy-consultations-and-reviews/advancing-new-zealands-energy-transition-consultation-document/introduction

Ministry of Business. (2024). Innovation and Employment: Electricity demand and generation data 1990–2024. https://www.mbie.govt.nz/building-and-energy/energy-and-natural-resources/energy-statistics-and-modelling/energy-statistics/electricity-statistics

Mirasgedis, S., Sarafidis, Y., Georgopoulou, E., Lalas, D. P., Moschovits, M., Karagiannis, F., & Papakonstantinou, D. (2006). Models for mid-term electricity demand forecasting incorporating weather influences. *Energy, 31*(2–3), 208–227. https://doi.org/10.1016/j.energy.2005.02.016

National Institute of Water and Atmospheric Research. (2024). Climate and weather data for New Zealand. https://data.niwa.co.nz/collections/climate.

Nowotarski, J., & Weron, R. (2018). Recent advances in electricity price forecasting. *Renewable and Sustainable Energy Reviews, 81*, 1548–1568. https://doi.org/10.1016/j.rser.2017.05.234

Pardo, A., Meneu, V., & Valor, E. (2002). Temperature and seasonality influences on spanish electricity load. *Energy Economics, 24*(1), 55–70. https://doi.org/10.1016/S0140-9883(01)00082-2

Paul, A., Padhy, S., & Dash, P. K. (2018). Sarima model based forecasting of household electricity consumption in india. *Journal of The Institution of Engineers (India): Series B, 99*(5), 577–583. https://doi.org/10.1007/s40710-017-0226-y

Shumway, R. H., & Stoffer, D. S. (2017). Time Series Analysis and Its Applications. New York: Springer. https://doi.org/10.1007/978-3-319-52452-8 4th edn.

Taylor, J. W., & Buizza, R. (2003). Using weather ensemble predictions in electricity demand forecasting. *International Journal of Forecasting, 19*(1), 57–70. https://doi.org/10.1016/S0169-2070(01)00123-6

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B, 58*(1), 267–288.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting, 30*(4), 1030–1081. https://doi.org/10.1016/j.ijforecast.2014.08.008

Ziel, F., & Weron, R. (2018). Day-ahead electricity price forecasting with high-dimensional structures. *Energy Economics, 70*, 396–420. https://doi.org/10.1016/j.eneco.2017.12.016

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B, 67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x