Published by
NIGERIAN SOCIETY OF PHYSICAL SCIENCES
Available online @ https://journal.nsps.org.ng/index.php/jnsps

J. Nig. Soc. Phys. Sci. 5 (2023) 1394

Journal of the Nigerian Society of Physical Sciences

Improving forecasting accuracy using quantile regression neural network combined with unrestricted mixed data sampling

Umaru Hassan[a,b], Mohd Tahir Ismail[a,*]

[a]School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Penang, Malaysia
[b]Department of Statistics, Federal Polytechnic Damaturu, 620221 Yobe, Nigeria

Abstract

A traditional regression method involving time series variables is often observed at the same frequencies. In a situation where the frequencies differ, the higher ones are averaged or aggregated to the lower frequency. A Mixed Data Sampling (MIDAS) regression model was introduced to address such problems. In any country, stakeholders are interested in monitoring and forecasting accurately the Gross Domestic Product (GDP) using the dynamics of macroeconomic variables. We applied the hybrid QRNN-U-MIDAS model to forecast quarterly GDP using monthly and weekly data. The Quantile Regression Neural Network (QRNN) is designed to model nonlinear relationships amongst data sampled at the same frequency. Therefore, we take advantage of QRNN skills using the optimization techniques of gradient descent-based algorithms to optimise the estimated loss function $E^a(\tau)$, and introduce them into the U-MIDAS framework, which can handle mixed data frequencies, and construct a QRNN-U-MIDAS model. The suggested hybrid QRNN-U-MIDAS model was implemented in an R-package that we created to perform both simulation and real-time data applications. The findings indicate that the QRNN-U-MIDAS regression model outperforms competing models in terms of its capacity for prediction across the conditional distribution of a response variable with a comprehensive view of the information contained in the variables, which is lacking in other competing models like U-MIDAS, ANN-U-MIDAS etc. Moreso, this novel model will add to the existing works of literature on robust forecasting models.

DOI:10.46481/jnsps.1394

Keywords: Neural network, Quantile regression, Unrestricted mixed data sampling frequency, Frequency alignments, GDP

Article History :
Received: 08 February 2023
Received in revised form: 14 June 2023
Accepted for publication: 05 July 2023
Published: 13 November 2023

Communicated by: T. Latunde

1. Introduction

The conventional forecasting models usually require the data to be sampled at the same frequency. Consider regression analysis as a statistical technique that serves as an investigating tool for the relationships among interrelated variables. Some econometric models have been developed to forecast time series data using explanatory variables; unfortunately, most macroeconomic indicators are not sampled at the same frequency. GDP data, for example, are sampled annually or quarterly, import and export values are sampled monthly, and most stock data are sampled daily [1]. Therefore, the idea of constructing regression models that combine data with different sampling frequencies emerged. This raises the issue of how to do empirical analysis on relationships between data collected at different frequencies.

Forecasting major economic factors like GDP growth is

*Corresponding author: Tel.: +60164143464;
Email address: m.tahir@usm.my (Mohd Tahir Ismail)

critical for both the central bank and a country's industry's decision-making processes. Economic data is impacted by the data sampling rate issue as well as the publishing delay. A preliminary estimate is needed because this data is valuable for creating and implementing policies. Andreou *et al*. [2] uses Mixed Data Sampling (MIDAS) to link variables sampled at various frequencies without missing high-frequency information. To achieve a parsimonious specification, MIDAS regressions are often based on distributed lag polynomials such as the exponential Almon lag [3]. In general, MIDAS regression is designed to bridge the gap between retaining the individual timing information of the high-frequency data and minimising the number of parameters to be estimated [1, 4].

Unrestricted MIDAS is a modified MIDAS that does not use any functional lag polynomials such as Almon or Beta. The use of U-MIDAS in macroeconomic applications is motivated by the fact that the difference between sample frequencies is not as large in many situations [5]. The Quantile regression neural network (QRNN) is flexible to represent linear relationships, including those with interactions between predictors, without prior specification of the form of the relationships by the modeler [6].

The superiority of the Quantile Regression Neural Network is that it reveals the entire conditional distribution of the dependent variables and also complex nonlinear problems, which can produce more accurate and informative results than other methods [6]. However, it can only model data of the same frequency. We introduce QRNN into the U-MIDAS framework work and construct the QRNN-U-MIDAS model. The proposed model can present an improved forecast with a comprehensive view of the information contained in the variables. Unlike MIDAS regression is often based on distributed lag polynomials such as the exponential Almon lag to achieve a parsimonious specification.

The remaining part of this paper continues from Section 2 with an extensive background of the study, followed by Section 3 detailed research methodology that provides the formulation of the model and analysis techniques. Section 4 presents both simulation and actual data results, and finally, Section 5 concludes the paper.

MIDAS models can use higher-frequency observations even when the corresponding lower-frequency data for the period is not yet available, and they can analyze time series data of different frequencies in the same regression without aggregation or interpolation [7, 8]. When the gap between frequencies is modest, and the risk of parameter proliferation is low [9], found that utilizing an unconstrained lag polynomial (U-MIDAS) can simplify model estimation while simultaneously improving forecasting performance. QRNN-MIDAS, which employs the restricted lag polynomial (e.g., exponential function), was recently proposed [10]. The class of MIDAS models uses the Lag Polynomials function to impose some weights on the regressors, which may cause the model's predictive power to decline if the imposed structure differs from the method used to generate the data [11]. All of these efforts were made to develop a better model to minimise the error in predicting macroeconomic variables like GDP, which depend on the outcomes of other variables observed at different frequencies. To this end, however, we developed a unique model we named QRNN-U-MIDAS by dropping the restrictions on the parameters that are caused by the use of Almon or Beta polynomials so that it can provide a better estimate while also increasing prediction performance along the conditional distributions of the response variable.

The benefits of the proposed model are that:

1. Unlike QRNN modeling, the suggested QRNN-U-MIDAS can directly use mixed frequency data (raw) without pre-processing.
2. By utilizing the QRNN framework, the QRNN-U-MIDAS model can detect nonlinear patterns and improve forecasts.
3. The QRNN-U-MIDAS is a novel model that proposes a method for bridging the gap between neural network-mixed data frequency analysis, such as ANN-U-MIDAS, MIDAS, and U-MIDAS. The distribution of the output variable using QRNN-U-MIDAS enhances the heterogeneous effect of the input variables, and detailed information is provided for decision makers.

The limitations of QRNN are:

1. The Neural Network is designed to model nonlinear, mixed relationships among data samples observed at the same frequencies and cannot dictate some valuable information in mixed data samples. As a result, we must use U-MIDAS to investigate the relationship between variables of various sample frequencies explicitly.
2. The fact that the error function of QRNN is not differentiable at some point, as its derivative is infinity at the origin, indicated that optimization methods of standard gradient-based ANN might prematurely converge, which leads to sub-optimal model parameters [12].
3. Model complexity can over fit the training data to noise rather than to signal when the number of input variables and hidden layers are large. To overcome this overfitting problem, some researchers proposed using weight decay regularisation and ensemble average via bootstrap aggregation Breiman [13] in He *et al*. [14]. QRNN can optimise the overall loss directly and can also train a single model for forecasting several quantiles [15].

In any country, stakeholders are interested in monitoring and forecasting accurately the Gross Domestic Product (GDP) using the dynamics of macroeconomic variables. We will utilize the novel QRNN-U-MIDAS model to estimate quarterly GDP using monthly and weekly financial variables of Japan and the United States. Looking at the complexity of some measurements in GDP, they're mostly taken every quarter of a year. As a result, weekly predictors such as claims and daily stock market returns, among others, are available while stakeholders await the publication of GDP for the next period. Based on this, high-frequency variables can be used to predict a lower-frequency variable.

For forecasting quarterly data, we will propose a collection of monthly macroeconomic indicators, providing data on industrial production, consumer price index, stock prices, etc., similar to other studies [9]. However, fewer variables could be considered by extracting significant factors utilizing principal component analysis (PCA). The principal features could then be applied as predictors to forecast GDP growth rates.The forecasting horizon will be generated from hm = 0, 1, and 2 quarters, as monthly steps h = 1/3, h = 2/3, etc. Root mean square errors (RMSEs) and other accuracy measurements are often used in calculating multi-step-ahead forecasts with the HF forecasting horizons of $hm = 1/m$, $2/m$, and m/m . This study also considers five quantiles: 0.1, 0.25, 0.5, 0.75, and 0.9.

2. Literature review

Determining how to investigate the diverse nonlinear relationships between variables on mixed data sampling frequency can be difficult. However, scholars have made an effort to offer some established approaches in an effort to address this difficulty. The Mixed Data Sampling regression put forward by Ghysels [4], among them which can handle the unprocessed mixed data sampling frequency by adding the functional lag polynomial weights. Other models to be discussed are, mixed data sampling frequency (MIDAS) and the Artificial Neural Network (ANN), and unrestricted mixed data sampling frequency methodologies.

However, these models were limited in their ability to address the issue fully as it relates to response variables conditional distribution across the quantiles. In order to examine the heterogeneous nonlinear interaction between factors on the frequency of the mixed data sample of the macroeconomic variables directly and thoroughly, we integrate the U-MIDAS method into the framework QRNN and create a hybrid quantile regression neural network with unrestricted mixed data sampling frequency, known as the QRNN-U-MIDAS model. The benefit of this method is that it shows both intricate nonlinear problems and the entire conditional distribution of the dependent variables.

Various novel models were proposed using both simulation and empirical studies about the extent to which financial variables can be used to predict economic activity. More specifically, gross domestic product (GDP), inflation rate, and so on have emerged with robust results, despite the brief literature. To forecast/nowcast macroeconomic variables, the mixed frequency data sampling frequency (MIDAS) regression was utilized alongside other promising models.

2.1. Mixed data sampling

Variables must be sampled at the same frequency in conventional time series regression methods. In the absence of this, higher-frequency data are averaged or aggregated into lower-frequency data, which results in the loss of some vital information [7]. According to Lements [8], MIDAS regression provides more accurate forecasts when the variables are not aggregated to the same frequencies, as reported in Franses [16]. Also, Armesto *et al.* [17] state that MIDAS regression accommodates variables observed at various frequencies. It presents a simple, flexible, and parsimonious class of time series models that allow the dependent and independent variables of time series regressions to be sampled at different frequencies.

MIDAS regression generally fills the gap between keeping the high-frequency data's individual timing information and minimising the number of parameters that must be estimated [1]. Because of the parametric restraints, the weight function is represented as a nonlinear parametric function in MIDAS models with minimal number of parameters [18, 19]. MIDAS regression models are popular because they exploit information in high-frequency data under a parsimonious setting. They handle issues of frequency mismatch between high and low frequencies without aggregating the data before model estimation by employing a weighting function that uses lag polynomials [20]. Most researchers used these three methods of processing mixed frequency data sampling for prediction.

Averaging the higher frequencies, as reported by Bams *et al.* [21] and Wang *et al.* [11], is one technique researchers use to use the mixed frequencies directly by applying the lag polynomial function to impose weights on the regressors or its hybrid where the model is not restricted. We call it the unrestricted mixed data sampling method developed by Baumeister *et al.* [22]. MIDAS models have been used in several studies to forecast quarterly time series utilising data collected monthly, weekly, or daily. Most recent research uses monthly or daily financial data to forecast quarterly time series. Since the critical performance indicators of forecasting's accuracy have confirmed the efficacy of the MIDAS-Almon model, researchers like Gunay *et al.* [23] and Götz & Hauzenberger [24] applied the MIDAS-Almon method to test the effect of the COVID-19 pandemic on some countries GDP, like the US, China, Indonesia, and so on. The model correctly predicts the decrease in the gross domestic product during these times.

Das *et al.* [25] and Babii *et al.* [26] presented a novel "mixed frequency-based regression approach," along with Functional Data Analysis (FDA), to study the impact of global unrest on stock market relationships. The output suggests that global crises generally affect global stock markets. The level of effect mainly depends on the nature and context of the crises that drive the feelings in financial markets. Various authors reported modelling with more than one predictor at different frequencies. For example. Khoo & Cheung [27] and Penev *et al.* [28]. MIDAS has provided researchers with new opportunities and possibilities to use any accessible data from various frequencies most efficiently in forecasting and nowcasting without having to deal with the issue of the varied lags of several macroeconomic time series variables [4]. Zhao *et al.* [29], introduced mixed data sampling regression models to forecast carbon dioxide emissions. Other researchers conducted a study on financial and macroeconomic problems such as stock market returns [30], GDP [31–34], and inflation [35]. Mixed frequency regression models provide a practical way to accept variables sampled at multiple frequencies, allowing for evaluating many specifications. Compared to naive benchmark models and estimators, different classes of these models produce a slight boost in out-of-sample prediction performance at close horizons [36].

2.2. Mixed data sampling frequency and other forecasting models

We will review some novel models that took advantage of the mixed data sampling frequency regression model to build robust estimation and forecasting models. Using penalized least-squares estimators and MIDAS to enforce smoothness via lag distribution Kapetanios [37] employed daily statistics to predict monthly inflation rates. According to the findings, the commodity price index (CPI) can be used to predict inflation rates [31]. To anticipate actual US GDP growth using crude oil prices, a time-varying parameter called (TVP-MIDAS) was adopted. The predictability of GDP growth varies depending on forecasting horizons. Using a regime-switching GARCH–MIDAS model, they looked into the relationship between oil price instability and its primary macroeconomic indicator. The authors claim that TVP-MIDAS beats the other models used in the study. The authors stated that GARCH-MIDAS models could significantly beat their single-regime counterparts when forecasting out-of-sample oil volatility. Building forecasting models that can provide a better prediction is encouraged by stakeholders to optimize operations and gain competitive advantages [38–40]. The unemployment rate is the best indicator to forecast quarterly GDP growth, according to a study by Kingnetr *et al*. [33]. The empirical findings showed that U-MIDAS outperformed MIDAS regardless of other indicators. Similar findings were reported in Foroni & Schumacher [5]. MIDAS-based models perform better than the traditional GARCH-based estimates and conditional quantile specifications, notably over multi-day forecast horizons [41]. The effect of "hot money" or money that moves quickly and often across financial markets and enables stakeholder lock-in at the best short-term interest rates, on the performance and instability of the Chinese stock market was investigated using the GARCH-MIDAS model.

The empirical findings reveal that there is no clear correlation between the hot money growth rate and the return on the Chinese stock market, meaning that hot money does not drive the Chinese stock market and vice versa [42]. However, since volatility is modelled as the combination of two variables, the Double Asymmetrical GARCH-MIDAS (DAGM) Model has the benefit of being more accurate [43]. Renato *et al*. [44] employed the MIDAS-quantile regression (MIDAS-QR) method proposed by Ghysels *et al*. [8] to explore the association among high-frequency predictor variables and low-frequency predictand variables at different quantiles. Alternatively, Almon's constraint may be excessively harsh on the underlying Data Generating Process (DGP). As a result, Foroni *et al*. [7] presented the unrestricted MIDAS model, which has no restrictions as a result of the lag polynomial weights. U-MIDAS [7] made an effort to compare unrestricted MIDAS (U-MIDAS) and MIDAS with distributed lag functions estimated by NLS and find out that U-MIDAS performs better in both simulated and empirical quarterly GDP and monthly predictors than MIDAS when sampling differences are small, similar to OLS, but not with large differences in sampling frequencies [45]. As the difference in frequency increases, the U-MIDAS becomes unappealing because of the parameter proliferation associated with high-frequency lag growth [46]. In summary, when the aggregated frequency is small, U-MIDAS regression outperforms the MIDAS model. Barsoum & Stankiewicz [9] investigated the utility of MIDAS models with unconstrained lag polynomials and a Markov-switching component for modeling massive datasets. For many macroeconomic applications, the unconstrained Markov-switching MIDAS model is an excellent choice for the constrained MS-MIDAS model, particularly when the frequency difference between the variables is small.

2.3. Model specifications

Using parameters of the lagged coefficients of B(k; θ) in a parsimonious fashion is one of the key MIDAS features. We call various specifications of MIDAS regression polynomials the" Exponential Almon Lag," a specification selection and adequacy testing. Besides the usual properties of the error term, other specifications of the MIDAS regression models need to be considered. We first select the functional constraints that will affect the model's precision and an appropriate maximum lag order should be chosen. The best way to address these issues is to use information criteria to select the appropriate model in terms of the parameter restriction and the lag orders using either in-sample or out-of-sample precision measures [47].

In order to estimate and interpret the UMIDAS model, key assumptions such as aggregation, constant coefficient, linearity, and uncorrelated measurement errors between high-frequency and low-frequency variables should be considered. Violations of these assumptions were addressed through data transformation, modifying model specifications, adding variables, using robust estimation techniques, running diagnostic tests, and exploring alternative models.

2.4. Nowcasting

Nowcasting, derived from the words "now" and "forecasting," has lately gained popularity in economics due to the rising demand for rapid, short-term economic analyses and projections. Data on crucial indicators, such as GDP and its components, is only released after a significant delay and is susceptible to modifications afterward [48, 49]. Short-term forecasting and nowcasting of economic activities are of interest for economic and policy decision-making during the Covid-19 crisis. Nowcasting is regarded as a crucial forecasting topic since it is frequently utilised as an input for models that are successful in the medium-term [50].

Using an extensive real-time dataset of about 550 macroeconomic indicators from New Zealand and worldwide, Richardson *et al*. [51] trains various popular machine learning algorithms to simulate a genuine nowcasting situation. The findings suggest that machine learning approaches improve nowcasting accuracy [20, 52]. The main premise of nowcasting is to construct an "early estimate" before the data is officially released by using information available earlier and at a higher frequency than the variable of interest [53].

2.5. Artificial Neural Network (ANN) and MIDAS

Other machine learning algorithms that might effectively model mixed sampling frequency are briefly mentioned in this section. Artificial neural networks (ANNs) are models based on biological neural networks and are used to approximate functions with a high number of inputs. They are typically depicted as systems of interconnected components that communicate with one another. Researchers have recently made a tremendous breakthrough in taking advantage of MIDAS regression to build numerous novel models that can provide better forecasting and dictate volatility in macroeconomic and financial data. Machine learning has not also been left behind in that regard. See, for example, a paper by Eds & Goebel [54], that discussed a hybrid of the mixed data sampling (MIDAS) regression and back propagation (BP) neural network (MIDAS-BP model) to forecast carbon dioxide emissions. The forecasting ability of MIDAS-BP is remarkably superior than MIDAS, OLS, polynomial distributed lags (PDL), and auto-regressive moving average (ARMA) models in terms of accuracy [29].

Neural network models have been successful in providing robust forecasting in many areas of research, such as demand forecast [55], energy market [15], CO_2 emissions [29] and several competing forecasts Pan *et al*. [31] in Challu *et al*. [55]. The popularity of neural network forecasting procedures is not restricted to industry but has also reached academia [56]. The versatility of an artificial neural network (ANN) in discovering nonlinear correlations or patterns among variables is well recognized [57]. Some of ANN's characteristics are self-learning, anti-jumping capacity, and data-driven [58].

Consequently, its robustness can be extended to nonlinear problems in areas such as economy [59], finance [11], energy management and environment [60], etc. A typical ANN model samples time series data at the same rate. In order to forecast Lahore, Pakistan's primary weather parameter, Artificial Neural Networking Multi-layer Perceptron (ANN-MLP) models were compared with the Exponential Smoothing Algorithm (ETS), and the Auto-Regressive Integrated Moving Average (ARIMA) models [38]. Application of ANN to macroeconomic variable forecasting has been demonstrated, for example, in a study by Galeshchuk [61], that predicted the exchange rates and that forecasted the Indian monthly inflation rate [36]. While Stevanović *et al*. [62], and Xu *et al*. [11] created a model called Artificial Neural Network Mixed Data Sampling (ANN-MIDAS) in recent years to investigate the nonlinear pattern contained in the variables, and the use of China's monthly inflation rate estimate serves as evidence of its efficacy.

Based on the QRNN-MIDAS method, frequency alignment is conducted on each high-frequency variables based on the maximum lag order as determined by information criteria. The frequency alignment is then given a weight function to create a low-frequency variable. This enables the QRNN model to deal with the MIDAS data in their raw form. In the electric power business, reliable and accurate load forecasting is essential for decision-making. Traditional point forecasting approaches are unable to account for uncertainty. This work suggests an improved QRNN (iQRNN) that leverages well-known deep-learning techniques. It is superior to regular QRNNs in terms of accuracy, stability, and computing efficiency [62].

However, these models were limited in their ability to address the issue fully as it's related to how the response variable was distributed along the quantiles. The QRNN model can simulate the nonlinear relationships between variables observed at the same frequencies. Still, it is unable to grasp the important information present in the mixed data sampling frequency. In contrast, U-MIDAS directly handles the raw mixed data sampling frequency without incorporating the polynomial lag weights. In order to instantly and thoroughly discover the heterogeneous relationship among the variables on the mixed data frequency of the macroeconomic variables, we integrate the U-MIDAS method into the QRNN framework and create a novel quantile regression neural network with unrestricted mixed data sampling frequency, known as the QRNN-U-MIDAS model.

The proposed QRNN-U-MIDAS model aids in preserving important information concealed in the real dataset by avoiding potential issues with frequency conversion. The ability to recognize complex nonlinear interactions between variables with heterogeneous sample frequencies is flexible. Additionally, it lessens the covariates' diverse impacts on the conditional distribution of a response variable. The advantage of the Quantile Regression Neural Network is that it displays the conditional distribution of the dependent variables and complex nonlinear issues, which can result in more accurate and detailed results than other approaches.

3. Methodology

We seek to develop a robust model that can display the entire conditional distribution of the dependent variable and complex nonlinear issues, which can result in more accurate and detailed results than other approaches for variables sampled at different frequencies. The theoretical basis and formulations of the models will be discussed. The novel model will be subjected to testing using macroeconomic variables and compared with other models to examine its efficacy. We consider QRNN alongside U-MIDAS approach to develop QRNN-U-MIDAS. The outcome QRNN-U-MIDAS will be considered for comparison among other competing models like ANN-U-MIDAS, U-MIDAS, and MIDAS.

Artificial Neural Networks (ANNs) are powerful in solving nonlinear problems and handling missing, noisy, and inconsistent data without requiring assumptions about data distribution. They learn from data and accept numeric inputs directly for mining purposes, making them applicable to a wide range of problems Xu *et al*. [11]. However, the performance of a neural network model can be influenced by factors like data quality, variable selection, and network architecture. For time series data, pre-processing steps such as outlier elimination, handling missing values, data normalization, and transformation were conducted before training a neural network.

For example, consider a three-layered QRNN with 'I' and 'h' input and hidden/inner layer neurons, respectively. The hyperbolic tangent sigmoid function will be used to activate the hidden layer, while the identity function will be used as the transfer function in the output layer. Frequency alignment will be conducted on the variables. An optimum solution procedure will be provided using the gradient-based optimization algorithm to renew biases and weights [11]. We wish to discuss the individual models separately.

3.1. U-MIDAS model

Similarly, because the U-MIDAS regression does not enforce any constraints on the parameters, no bias can arise due to a potentially inaccurate restriction on coefficients. As a result, when k is just relatively smaller, the U-MIDAS regression model provides a good choice for modeling and forecasting [7]. Suppose low-frequency data is measured quarterly while high-frequency data is measured monthly. The U-MIDAS model for h-step forecasting can be specified as:

$$y_t^Q = \beta_0 + \sum_{j=1}^{k} \beta_j x^m_{t-h-(j-1)/m} + \varepsilon_t, \quad (1)$$

where y_t^Q is a quarterly response variable, $x^m_{t-h-(j-1)/m}$ as monthly predictors considered at $j-1$ months before the last month $t-h$ of the quarter, h represents the forecasting horizon, frequency mismatch m, which is 3 in this case, and k being the number of predictors used to predict y_t^Q.

For the purpose of demonstration, there are three months in a quarter, $m = 3$, let $k = 4$ and $h = 2$, the equation will be as:

$$\begin{aligned} y_t^Q &= \beta_0 + \beta_1 x^3_{t-2} + \beta_2 x^3_{t-2-(2-1)/3} + \beta_3 x^3_{t-2-(3-1)/3} + \beta_4 x^3_{t-2-(4-1)/3} \\ &= \beta_0 + \beta_1 x^3_{t-2} + \beta_2 x^3_{t-2-1/3} + \beta_3 x^3_{t-2-2/3} + \beta_4 x^3_{t-3}. \quad (2) \end{aligned}$$

This implies if y_t^Q is the growth rate for the first quarter of 2020, x^3_{t-2} refers to the value of the predictor from September 2019. While x^3_{t-3} refers to the values from June 2019 and so on. However, the best specifications for MIDAS and U-MIDAS models are based on the information criterion (AIC, BIC, or GACV).

The common identification issues in U-MIDAS regressions include overparameterization, misspecification of the functional form, data availability challenges, and weak instruments, all of which can hinder the accurate estimation and interpretation of parameters. Researchers focus on proper model specification, variable selection, and robust estimation techniques tailored to the specific data context in U-MIDAS regressions to address these issues. Several approaches can be employed to detect overparameterization in U-MIDAS models: model selection criteria like AIC or BIC, or GACV, assessing variable importance and significance, conducting sensitivity analysis, diagnosing collinearity, and considering expert knowledge and economic theory. These techniques help determine if the model has excessive parameters that may not be necessary for accurately describing the relationship between variables.

Using QRNN-U-MIDAS regressions with macroeconomic and high-frequency financial data, identification issues such as overfitting due to model complexity and the choice of neural network architecture, as well as misspecified lag structures and input variables, can impact the reliability and generalization of the model's predictions. Proper model selection, regularization techniques, and careful consideration of lag structures and input variables are essential for addressing these identification challenges, leading to more meaningful predictions of GDP and inflation. However, we recommended for further studies to take care of regularization techniques.

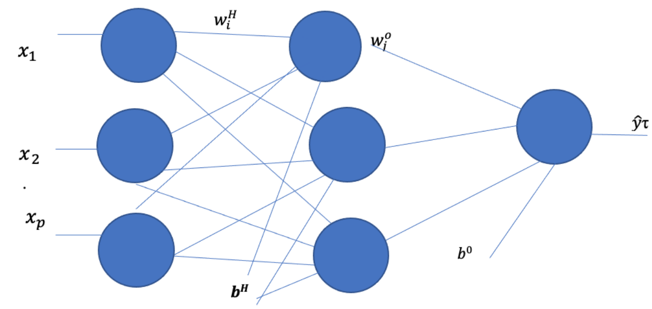

Figure 1. Schematic QRNN model neural network.

3.2. QRNN model

Suppose a forecasting model consisting of $x_1, x_2, x_3, \ldots, x_p$ predictors, corresponding to 'p' input neurons, which connect to 'n' hidden neurons of a single layer and connect to one output neuron to give out the prediction. We present Figure 1 as the schematic architecture of neural networks for the purpose of demonstration.

The j-th hidden layer node output is

$$g_{j,t} = \tanh(\sum_{i=1}^{p} x_{it} w_{ij}^H + b^H), \quad (3)$$

where w_{ij}^H is the weights connecting the i^{th} node of the input to j_{th} node of the hidden layer $j = 1, 2, \ldots, n$ and b^H is the bias to the j^{th} node.

The estimate $\widehat{y_t^\tau}$ (conditional (tau) τ- quantile) is given by;

$$\widehat{y_t^\tau} = f^o \sum_{i=1}^{p} g_{j,t} w_j^o + b^o, \quad (4)$$

where $\widehat{y_t^\tau}$ = estimated conditional τ- quantile.
f^o = output transfer function.
w_j^o = weight connecting the output node.
b^o = bias of the output.

Usually, Back Propagation (BP) algorithm is used to determine weights and biases in QRNN at different quantiles. However, it has a drawback in computational time and delays in convergence rate, among other things. The Particles Swarm Optimization (PSO) is one of the ways to overcome these drawbacks [6]. QRNN is flexible to represent linear relationships, including those with interactions amongst variables without earlier design by Pradeepkumar & Ravi [6]. The superiority of the Quantile Regression Neural Network is that it reveals the entire conditional distribution of the dependent variables and also complex nonlinear problems, which can produce more accurate and informative results than other methods [6].

3.3. Building QRNN-U-MIDAS model

Given $(y_t)_{t=1}^{N}$ as the response variables observed at low frequency (LF), and a high frequency (HF) predictor as $(x_t)_{i=1}^{I}$, m_i is the frequency associated with the variable i which is m_i times higher than the lower frequency variable. This is sometimes called frequency mismatches between $(y_t)N$ and$(x_t)_{i=1}^{I}$. Also, $(L_i)_{i=1}^{I}$ is the maximum lag order. The following steps are used to build the QRNN-U-MIDAS model estimation.

1. Frequency alignment is conducted on the input variable x_{ti} to get $x_{i,t-hi}$, $x_{i,t-1/mi-hi}$, $x_{i,t-2/mi-hi}$, . . . , $x_{i,t-Li/mi-hi}$, so that x_{ti} is transformed into low frequency. By so doing, therefore, the MIDAS variables are transformed to be of equal frequencies as the response variable. Consequently, it can be solved by ordinary least squares (OLS), and at this stage, it can be introduced into a neural network [63].
2. The jth hidden layer node $g_j(\tau)$ is considered by introducing a sigmoid transfer function $f^{(H)}$ to give:

$$g_j(\tau) = f^{(H)}\left(\sum_{i=1}^{I} w_{ij}^{H}(\tau)\, x_{i,t-hi} + b_j^{H}(\tau)\right), \quad (5)$$

where $w^H(\tau)$ =($w_{11}^H(\tau)$, $w_{12}^H(\tau)$, ..., $w_{ij}^H(\tau)$, $w_{21}^H(\tau)$, ... , $w_{IJ}^H(\tau))^T$ is considered as hidden layer weights, $b_j^H(\tau)$ is the vectors of the hidden layer bias, $f^{(H)}$ is a sigmoid transfer function applied to inner product between the hidden layer weight and predictors, plus the hidden layer bias using hyperbolic tangent function.

$$f^{(H)} = \frac{\mathbf{e}^{x} - \mathbf{e}^{-x}}{\mathbf{e}^{x} + \mathbf{e}^{-x}}. \quad (6)$$

3. Consequently, the τ^{th} conditional quantile $\hat{y}(\tau)^{th}$ is given as;

$$\hat{y}(\tau) = f^{(o)}\left(\sum_{i=1}^{I} w_i^{o}(\tau)\, g_j(\tau) + b_j^{o}(\tau)\right), \quad (7)$$

where $w^o(\tau) = (w_1^o(\tau), w_2^o(\tau), . . ., w_j^o(\tau))^T$ is a weight associated with the output layer, $b_j^o(\tau)$ is the output bias, and $f^{(o)}$ transfer function selected based on the task, using identity function.

Adding up the three layers of the model we get similar to what is obtained in Xu *et al.* [10].

$$\widehat{y_t^{\tau}} = f^{(o)}\Big(\sum_{i=1}^{I} w_i^{o}(\tau)\, f^{(H)}\Big(\sum_{i=1}^{I}\sum_{I}^{L_i} w_{iIj}^{H}(\tau)\, x_{i,t-1/mi-hi} + b_j^{H}(\tau)\Big) + b^{o}(\tau)\Big), \quad (8)$$

where all the parameters are as defined above. Usually, the Back Propagation (BP) algorithm is used to recompute weights and biases in neural networks at different quantiles lays for optimization.

Figure 2 shows how the input variables are introduced into the neural network and how frequency alignments are handled at the inner layer. Both feed-forward and back propagation are indicated.

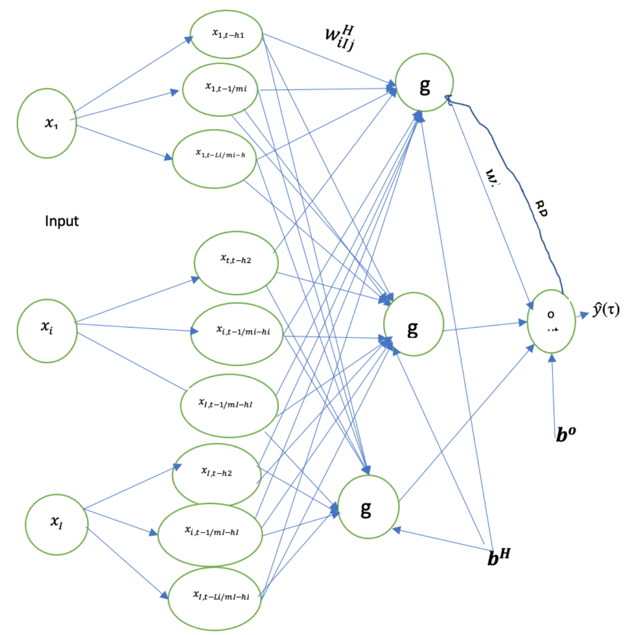

Figure 2. Feed forward/feed backward QRNN-U-MIDAS with three layers showing frequency alignment.

3.4. Error function

To estimate the error function of the QRNN-U-MIDAS model, which we seek to minimize, we define as

$$E = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}t)^2. \quad (9)$$

The Huber norm is used to build approximations of $\rho^{\tau}u$.

$$\rho^{\tau}u = \begin{cases} \tau h(u), & if\ u \geq 0 \\ (\tau - 1)\, h(u), & if\ u < 0 \end{cases}, \quad (10)$$

h(u) as the Huber function.

$$\text{h(u)} = \begin{cases} \frac{u^2}{2\varepsilon} & if\ 0 \leq IuI \leq \varepsilon \\ IuI - \frac{\varepsilon}{2}, & if\ IuI > \varepsilon \end{cases}.$$

h(u) provides a shift between squared and absolute errors about the origin.

$$E^{\tau} = \frac{1}{N}\sum_{t=1}^{N}\rho^{\tau}(y_t - \hat{y}t), \quad (11)$$

where N denotes the number of observations, y_t *is* the observed value, $\hat{y}_t$ is the predicted output value. We adopt and use the Huber norm to create smooth approximations of ρτ(u) [11]. We presented the Huber function as $h(u)$ and the given magnitude threshold as ε.

The QRNN-U-MIDAS loss E(τ) is revised using the approximate check function ρ_{τ}^{a} (u). The approximate loss function in this case is:

$$\mathrm{E^a(\tau)} = \frac{1}{\mathrm{N}}\sum_{\mathrm{t=1}}^{\mathrm{N}} \rho^{a}\tau(\mathrm{y_t} - \widehat{\mathrm{y}}_{\mathrm{t}}^{\tau}) \quad (12)$$

This approximate loss $E^a(\tau)$ is always greater than the curve's differentiability ρ^a_τ (u), which smoothes out at any value of τ. Using the optimization techniques of gradient descent-based algorithms to optimise the estimated loss function $E^a(\tau)$. Using the partial derivatives with respect to the parameters by the chain rule method, which is presented step by step below.

1. The derivative of $E^a(\tau)$ with respect to output weights $w^{(O)}(\tau)$.
$$\frac{\partial E^{(a)}_\tau}{\partial w^{(o)}_\tau} = \frac{\partial E^{(a)}_\tau}{\partial \widehat{y}^{(\tau)}_t} X \frac{\partial \widehat{y}^{(\tau)}_t}{\partial w^{(o)}_\tau} = -\frac{1}{T}\rho^a\tau(y_t - \widehat{y}^\tau_t).\ g^{(\tau)}, \quad (13)$$
where $g^\tau = ((g^\tau_1, g^\tau_2, g^\tau_3, \ldots, g^\tau_j))^T$,
$$\rho^a\tau(y_t-\widehat{y}^\tau_t).\ g^{(\tau)} = \begin{Bmatrix} \tau, & (y_t-\widehat{y}^\tau_t)>\varepsilon \\ \tau(y_t-\widehat{y}^\tau_t)/\varepsilon, & 0\le(y_t-\widehat{y}^\tau_t)\le\varepsilon \\ (1-\tau)(y_t-\widehat{y}^\tau_t)/\varepsilon, & -\varepsilon\le(y_t-\widehat{y}^\tau_t)<0 \\ -(1-\tau), & (y_t-\widehat{y}^\tau_t)<-\varepsilon \end{Bmatrix}.$$
2. The derivative for the hidden layer weight vector $w^{(H)}_\tau$ is obtained as:
$\frac{\partial E^{(a)}_\tau}{\partial w^{(H)}_\tau} = \frac{\partial E^{(a)}_\tau}{\partial \widehat{y}^{(\tau)}_t} X \frac{\partial \widehat{y}^{(\tau)}_t}{\partial g^{(\tau)}} X \frac{\partial g^{(\tau)}}{\partial w^{(H)}_\tau} = -\frac{1}{N}\ \rho'^a\tau(y_t-\widehat{y}^\tau_t).\ w^{(o)}_\tau.\ f'^H((w^{(H)}_\tau)^T x_{t-hi} + b^{(H)}_j\tau).\ x_{t-hi}$,
where $x_{t-hi} = (x_{1,\ t-hi},\ x_{2,\ t-hi}, \ldots, x_{I,t-hi})^T$, $f'^H((w^{(H)}_\tau)^T x_{t-hi} + b^{(H)}_j\tau)$.
3. The derivative for the output layer biases $b^{(o)}(\tau)$ is obtained as:
$$\frac{\partial E^{(a)}_\tau}{\partial b^{(o)}_\tau} = \frac{\partial E^{(a)}_\tau}{\partial \hat{y}^{(\tau)}_t} X \frac{\partial \hat{y}^{(\tau)}_t}{\partial b^{(o)}_\tau} = -\frac{1}{N}\rho'^a\tau(y_t - \hat{y}^\tau_t). \quad (14)$$
4. The derivative for the hidden layer biases $b^{(H)}(\tau)$ is obtained as:
$$\frac{\partial E^{(a)}_\tau}{\partial b^{(H)}_\tau} = \frac{\partial E^{(a)}_\tau}{\partial \hat{y}^{(\tau)}_t} X \frac{\partial \hat{y}^{(\tau)}_t}{\partial g^{(\tau)}} X \frac{\partial g^{(\tau)}}{\partial b^{(H)}_\tau} = -\frac{1}{N}\rho'^a\tau(y_t - \hat{y}^\tau_t).\ w^{(o)}_\tau.\ f'^H((w^{(H)}_\tau)^T x_{t-hi} + b^{(H)}_j\tau). \quad (15)$$

In summary, the process of the optimization procedure is as follows:

The initial weights and biases ($w^{(o)}_{io}$, $w^{(H)}_{io}$ $b^{(o)}_{io}$ and $b^{(H)}_{io}$) are estimated randomly. The next sets of parameters as an output of the derivatives are ($w^{(o)}_{i1}$, $w^{(H)}_{i1}$ $b^{(o)}_{i1}$ and $b^{(H)}_{i1}$) this iterative process will continue until the optimal loss values is obtained. The maximum iteration K and threshold magnitude ε $= (2^{-8}, 2^{-9}, \ldots, 2^{-32})^T$ are set to propagate the loss function $E^a(\tau)$ back to all neurons. Each time the derivatives are calculated, the parameters are updated. This is the fundamental idea of backpropagation, which is frequently used to train neural networks and is effective at resolving this minimization issue. Alternatively put, the approximation check function $\rho^{(a)}_\tau(u)$ gets closer to the initial check function $\rho^\tau(u)$ ie, $\lim_{\varepsilon\to 0} \rho^{(a)}_\tau(u) = \rho^\tau(u)$. In a similar vein, we also have
$\lim_{\varepsilon\to 0} E^a(\tau) = \lim_{\varepsilon\to 0} \frac{1}{N}\sum^N_{t=1} \rho^a\tau(y_t-\widehat{y}^\tau_t) = \frac{1}{N}\sum^N_{t=1} \rho_\tau\ (y_t-\widehat{y}^\tau_t = E(\tau)$.

Collectively, as ε approaches 0, the entire optimization process comes to a conclusion and converges to the least E(τ).

3.5. Forecasting design

In most cases, the variables that occur most frequently are presented at $N + d/m_i$, where, d indicates the most frequent observations that start the release of $y_t(\tau)$, as $d = 0, 1, 2, \ldots, m_i$. We observe the forecast of $y_t(\tau)$ based on the multi-step-ahead technique. We consider this as the distribution of quantiles yt+h(τ). From here, we involve quantile operators on QRNN-U-MIDAS model in Eq. (8) from both sides to obtain:
$$E(\hat{y}_t\ (\tau; h/\Omega N, d) = f^o \sum^J_{j=1} w^{(o)}_j(\tau)\ f^H \sum^I_{i=1}\sum^L_{l=0} w^{(H)}_{ilj}(\tau) E(x_{i,\ t+h-\ \frac{l}{mi-hi}}/\Omega N, d) + b^{(H)}_j(\tau)) + b^{(o)}_j(\tau). \quad (16)$$
$\Omega_{t,d} \equiv \left\{yt(\tau)^N_{t=1}, (x^N_{t1})^{N+d/m1}_{t1/m1}, \cdots, (x^N_{tI})^{N+d/mI}_{tI/mI}\right\}$, and $m = \max(m_1, m_2, m_3 \cdots, m_I)$,
where
$$\phi y_{t+h}(\tau)\,|\Omega_{t,d} = \begin{cases} \hat{y}_t\ (\tau; h), & h > 0 \\ y_t\ (\tau; h), & h \le 0 \text{ otherwise.} \end{cases}$$
h has a horizon of forecast on the release of the low-frequency predictor $(yt)^N_{t=1}$. Then, forecasts of the QRNN-U-MIDAS model as:
$$\hat{y}_t\ (\tau; h) = f^o \left\{\sum^J_{j=1} w^{(o)}_j(\tau)\ f^H \left\{\sum^I_{i=1} w^{(H)}_j(\tau) x_{i,\ t+h-\ \frac{l}{mi}\ -hi} + b^{(H)}_j(\tau)\right\} + b^{(o)}_j(\tau)\right\}. \quad (17)$$
It is important to note that the multistep-ahead forecast involves a link amongst the low and the higher frequencies forecast horizons h_i, ie., $h = [h_1] = [h_1] = \cdots = [h_1]$ where [.] is the upper limit function.

3.6. Model selection strategy

Determining the maximum lag order values for the QRNN-U-MIDAS model estimation is essential. $L \equiv (l_1, l_2, l_3, \cdots, l_I)^T$ and the number of nodes 'J' in the hidden layer. Using the generalized approximate cross-validation (GACV) criterion suggested by Xu *et al.* [10], we select the ideal pairing of L and J and describe it as:
$$GACV(L, J; \tau) = \frac{1}{N - df}\sum^N_{t=1} \rho_\tau\ (y_t-\widehat{y}^\tau_t), \quad (18)$$
where df is a measure of the fitted model's effectiveness in terms of dimensions, and it may be calculated by $\sum^N_{t=1} Cov(\hat{y}^{(\tau)}_t, y_t)/\sigma^2 y_t$.

We adopted GACV because, comparisons of GACV, the Akaike information criterion (AIC), and the Schwarz information criterion (SIC) demonstrate that GACV surpasses the other two in terms of accuracy and computing ease [64]. The GACV criterion also helps to drastically lower computing complexity in Xu *et al*. [10]. The ideal parameter combination (L,J) in this investigation are determined by:

$$(L^{+}, J^{+}) = arg\min_{L,J} (\text{GACV (L, J)}). \quad (19)$$

The best combinations of (L^+,J^+) with the lowest GACV parameters are chosen at the end.

3.7. Simulations design

Here we consider Monte Carlo simulations to demonstrate our novel model's ability to make better predictions. We will compare the QRNN-U-MIDAS model with other standard U-MIDAS, MIDAS and ANN-U-MIDAS models. The following stages are being used to design the process.

Stage 1: We use low-frequency variable yt of size N = 250, which is equivalent to making 250 quarterly observations. High-frequency variables, $x_{i,t-l/m1}$, and $z_{i,\ t-l/m2}$. Where x_i is the monthly variable while z_i refers to the weekly variables in this study.

Stage 2: In the trials, we use a rolling forecasting method to assess the performance of the suggested model in forecasts made more than one step in advance. The first estimation window has a fixed size of N_1 and ranges from 1 to N_1. In order to move along the entire data set for every rolling window, the estimating period includes the newest observation while tossing out the oldest. The following representative quantiles 0.1, 0.25, 0.5, 0.75, and 0.9 - are presented along with the optimal lag order values and the inner layer nodes, which are determined by the GACV.Taking into consideration the choice frequencies of two Variables, where, L_1=2,3,4,··· ,9, L_2=2,3,4,··· ,12 and J = 6,7,8,,···16, then each pair (L,J) is used to train the model. Thus, choosing the ideal pair (L,J) for training models is made simple by using a data-driven approach like GACV [65].

Table 1 only contains the parameter optimal values in order to conserve space. The neural network will use the best lag order and the appropriate number of hidden nodes to train the model.

We additionally take into account three distinct high frequency prediction horizons $h_i = 1/m, m/m$, and $2m/m$ and the equivalent low frequency forecast horizons $h = 0, 1, 2$in order to demonstrate the resilience of the QRNN-U-MIDAS model.

Stage 3: We determine the RMSE (root mean squared error) and MAPE (mean absolute percentage error) values to evaluate the models' predictive power.

They are defined as:

$$RMSE(\tau) = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(y_t - \widehat{y}_t^{\tau})^2}, \quad (20)$$

$$MAPE(\tau) = \frac{1}{N}\sum_{t=1}^{N}|(y_t - \widehat{y}_t^{\tau})| * 100. \quad (21)$$

The performance metrics values of RMSE and MAPE are smaller.

Stage 4: We run the aforementioned simulation 500 times, and then we present the results along with the Diebold Mariano (DM) test for the accuracy metrics. The test that Diebold & Mariano [66] first suggested takes into account a sample path of loss differentials $\{dt\}_{t=1}^{N}$. For a squared loss function, the formula is $dt = e_t^2 - e_t'^2$. The sample average, $\overline{d}$ asymptotically approaches a normal distribution under the presumption that the loss differential is a covariance stationary series:

$$\sqrt{N}\overline{d} \xrightarrow{d} N(\mu, 2\pi f_d(o)). \quad (22)$$

In particular, they proposed to test the null hypothesis that the forecast errors coming from the two forecasts bring about the same loss: $E[e_t^2 - e_t'^2]=0$ against the two-sided alternative. Thus, the resulting p-values represent the probability of obtaining the realized forecast error differential or a more extreme one in a new experiment if the null hypothesis was actually true. e_t^2= forecast error; $e_t'^2$ = benchmark error; dt= differential loss.

The test-statistic that will be used to calculate our p-values is computed as follows:

$$DM = \frac{\overline{d}}{\sqrt{\frac{2\pi \hat{f} d(o) N}{N}}}, \quad (23)$$

where, $\overline{d}=\sum_{t=1}^{N} dt$, $2\pi \hat{f} d(o)$ is a consistent estimate. Consider $2\pi \hat{f} d(0) = \sum_{\tau=-N-1}^{N-1} w_\tau \gamma d(\tau)$, where $\gamma d(\tau) = \frac{1}{N}\sum_{t=|t|+1}^{N}(dt - \bar{d})(dt - |\tau| - \bar{d})$.

3.8. Data generating process (DGP)

In designing our simulation, we consider some important features like mixed-frequency variables. Two independent predictors, x and z, are normally distributed ($i = 1, 2, \cdots, N$), with a monthly to quarterly frequency mismatch of 3 and a weekly to quarterly frequency mismatch of 12. We now use these higher frequencies to generate the lower frequency response variable (quarterly in this case). The next step is to design a pattern of relationship between the dependent and independent variables with heteroscedastic error terms. Finally, we develop our data generation process as equation (20) with some hyperparameters.

$$y_t = \text{trend} + \sin\left[w_1\left(\sum_{i}^{l1}\beta_i x_{1,t-l1/m1}\right)\right] + \exp\left[w_2\left(\sum_{i}^{l2}\lambda_i (z_{i,t-l2/m2})^2\right)\right] + \sigma_x \varepsilon t, \quad (24)$$

where $\varepsilon_t \sim$ N(0,1), , $\sigma_{x,}$= 0.01, σ_z = 001, for $t = 1, 2, \cdots, N$ and N = 500, $m_1 = 3, m_2 = 12, l_1 = 2, l_2 = 6, w_1 = 1, w_2 = 0.5$, trend = 1.5.

4. Results and discussion

4.1. Simulation results

This part presents the Monte Carlo simulations to show the goodness-of-fit and accuracy benefits of our hybrid QRNN-U-MIDAS model. We contrast the standard U-MIDAS, MIDAS, and ANN-U-MIDAS models with the QRNN-U-MIDAS model. Below are the findings of the simulation.

Table 1. The optimal parameters selected by the GACV for the models across quantiles.

		MIDAS		U-MIDAS		ANN-U-MIDAS		QRNN-U-MIDAS	
Error	τ	(L_1,L_2)	GACV	(L_1,L_2)	GACV	(L_1,L_2,J)	GACV	((L_1,L_2,J)	GACV
N(0,1)	0.10	(6,10)	0.290	(6,12)	0.276	(2,10,16)	0.246	(2,4,16)	0.246
	0.25	(4,10)	0.505	(6,12)	0.567	(4,12,16)	0.533	(6,12,14)	0.533
	0.50	(4,12)	0.461	(4,5)	0.486	(4,8,14)	0.473	(6,8,14)	0.433
	0.75	(6,12)	0.442	(2,10)	0.364	(2,10,16)	0.425	(2,10,16)	0.425
	0.90	(6,10)	0.258	(2,10)	0.267	(2,8,12)	0.312	(2,8,15)	0.312

NB: Using a normal distribution the following lags and nodes values were selected for each competing models.

The R package QRNNUMIDASV3 was developed and it depends on other R packages ("midasr", "qrnn", & "nnet"):

The equation below presents the estimation of the yt series with a varying lag order between 0 and 7 for variables xt and between 0 and 16 for variables zt. The results are shown in Table 2, where the parameters of the lag orders and their significance probability (p-values) are shown with the descriptive statistics in the first row. The process converged after 354 iterations, and only the significant parts were presented for the purpose of space.

```
>dt1=list(yt=yt,xt=xt,zt=zt)
>
>compare_models2(dt=dt1,k1=7,k2=16,m1=3,m2=12,quantile
=c(0.1,0.25,0.5,0.75,0.9),hidden=3,max.iteration=500)
```

The process in Table 2 converged at 354 iterations.
Resid. standard error: 0.9383 on 221 degree of freedom
(1 missingness observation deleted)
Multiple R-squared: 0.9754, Adjusted R-squared: 0.9635
F-statistic: 578.3 on 26 and 221 df, P-value: $<$ 2.2e-16

With R-square of 0.9654, we can comfortably claim that the predictor variables have explained up to 97% of the variation in yt series.

Tables 3, 4 and 5 provide the accuracy measured across the conditional distributed quantiles of the simulated data application of Japan. The first column presents the competing models, and the metric columns shows the error measures with the minimum error in bold figures. That is, the smaller the better in terms of predictions.It is also obvious that the QRNN-U-MIDAS model has the least error based on the simulated data when compared to other sibling models like U-MIDAS, MIDAS, and ANN-U-MIDAS models.

4.2. *Empirical application*

4.2.1. *Data*

The success of the QRNN-U-MIDAS model for empirical data is further evaluated in this section. We employ the hybrid QRNN-U-MIDAS model specifically to predict quarterly GDP growth in Japan and the US to see the efficacy of the model and compare its performance to that of the U-MIDAS, MIDAS, and ANN-U-MIDAS models. In terms of prospective low- and high-frequency variables to consider, we looked at the following variables to be the most effective in determining GDP growth and also used for forecasting or nowcasting.

Table 6 provides the list of most prominent variables used in predicting gross domestic product of any nation. The selection process aims to find the best subset of predictors. To guarantee the data's stationarity, we transform the original variables. The vast majority of these became stationary at the log of the first difference. Monthly data, such as the industrial output volume index, consumer price index, Real M2 money stock, and so on, can be used to estimate GDP [53]. The GACV, BIC, or AIC will be used to choose the best values for the lag order and the inner layer nodes. We generate three nowcasts for monthly horizons h_m = 0, 1, 2, and GDP growth estimates for the next three quarters (monthly horizons h_m = 3, 6, 9). For its implementation, we developed an R package called QRNNUMIDASV3, which depends on some R packages as (*"Midasr", "qrnn", "nnet"*). For the purpose of testing the package, we used data from two different countries, namely Japan and the United States.

4.3. *Findings*

Gross domestic product (GDP) is a key macroeconomic goal in Japan, and the government takes it seriously, monitoring its dynamics as it does in every other country. Without a doubt, precise real-time GDP estimation and forecasting can aid policymakers' understanding of the nature of growth and aid in creating useful monetary policy tools. Many times, a wide range of predictors, including macroeconomic indicators, is required for an accurate projection of GDP.

In this paper, we employ the proposed QRNN-U-MIDAS model to timely and precise GDP prediction. Monthly and weekly macroeconomic indicators are created by breaking down the predictions. The predictors, including GDP, real M2, IPI, houses, stocks, S&P500, etc., can be obtained from the "Genius finance database FRED (Federal Reserve Economic Database, https://fred.stlouisfed.org/)". We achieve stationarity at the first difference for most of the variables, except for the national housing index and the weekly common stock index at second differencing.

Figure 3 presents the time series of GDP, and Table 7 provides summary statistics for all variables from Jan. 1, 1994, to Dec. 1, 2021, comprising mostly monthly observations and one weekly observation.

From Figure 3, GDP is available as from 1994, quarter 1, until 2022, first quarter, with roughly 113 quarterly indicators within the range of time (depends on the publication delay). The core source of the data is FRED (Federal Reserve Economic Database, https://fred.stlouisfed.org/).

Figure 4 depicts the forecast for the next 12 quarters, or three-year periods, of GDP growth, along with the 95% confi-

Table 2. Showing the parameter estimate across the different lag orders for the variables and their respective p-values.

Min '= -2.2551	Q1 = -0.6389	Median = 0.1073	Q3 = 0.6680	Max = 2.7807
Coefficients:				
Estimates	std	error	t-value	pr(>|t|)
Intercept	1.9694327	0.1261210	15.615	< 2e-14 ***
Trends	0.1000072	0.0008679	114.047	< 2e-14 ***
mls(x, k = 0:7, m = 4)x.0/m	0.4267124	0.0643322	8.189	2.07e-12 ***
mls(x, k = 0:7, m = 4)x.1/m	0.3683006	0.0641479	5.896	1.37e-07 ***
mls(x, k = 0:7, m = 4)x.2/m	0.2869682	0.0680465	2.762	0.006219 **
mls(x, k = 0:7, m = 4)x.3/m	-0.004230	0.0658730	-0.080	0.936658
-	-	-	-	-
-	-	-	-	-
mls(x, k = 0:16, m = 12)z.0/m	0.2661054	0.0618960	5.931	1.14e-07 ***
mls(x, k = 0:16, m = 12)z.1/m	0.3402301	0.0599615	5.841	1.82e-07 ***
mls(x, k = 0:16, m = 12)z.2/m	0.4414556	0.0659546	6.845	7.81e-10 ***
mls(x, k = 0:16, m = 12)z.3/m	0.2633647	0.0577702	6.463	6.41e-8 ***
mls(x, k = 0:16, m = 12)z.4/m	0.2603666	0.0700954	5.150	5.75e-06 ***
mls(x, k = 0:16, m = 12)z.5/m	0.1254744	0.0632036	3.411	0.00660 ***
mls(x, k = 0:16, m = 12)z.6/m	0.067153	0.0630194	1.029	0.324821
mls(x, k = 0:16, m = 12)z.7/m	0.0566571	0.0673284	0.989	0.313954

signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1

Table 3. Summary of the accuracy measures from the simulated data τ=0.1 and 0.25.

Model	Performance measures at 0.1				Performance measures at 0.25			
	MSEP	RMSEP	MAEP	MAPE	MSEP	RMSEP	MAEP	MAPE
U-MIDAS	**0.4810**	**0.0694**	0.0537	0.0137	0.4710	0.0694	0.0517	0.0127
MIDAS	0.9002	0.0949	0.0740	0.0182	0.9002	0.0949	0.0740	0.0182
QRNN-U-MIDAS	0.6593	0.0812	**0.0421**	**0.0108**	**0.2084**	**0.0457**	**0.0225**	**0.0051**
ANN-U-MIDAS	33.6763	0.5803	0.4912	0.1150	31.9278	0.5650	0.4761	0.1104

NB: The first column presents the competing models while the second row presents the metric for 10% and 25% quantiles

Table 4. Summary of the accuracy measures from the simulated data τ=0.5 and 0.75.

Model	Performance measures at 0.5				Performance measures at 0.75			
	MSEP	RMSEP	MAEP	MAPE	MSEP	RMSEP	MAEP	MAPE
U-MIDAS	0.4709	0.0604	**0.0217**	0.0125	0.4601	0.0694	0.0427	0.0227
MIDAS	0.9012	0.0922	0.0520	0.0170	0.7011	0.0915	0.0731	0.0172
QRNN-U-MIDAS	**0.2810**	**0.0530**	0.0237	**0.0055**	**0.4239**	**0.0651**	**0.0317**	**0.0084**
ANN-U-MIDAS	33.6763	0.5803	0.4900	0.1120	30.2015	0.5496	0.4514	0.1069

NB: The first column presents the competing models while the second row presents the metric for 50% and 75% quantiles

Table 5. Summary of the accuracy measures from the simulated data τ = 0.9.

Model	Performance measures at 0.9			
	MSEP	RMSEP	MAEP	MAPE
U-MIDAS	0.4905	0.0694	0.0512	0.0128
MIDAS	0.7202	0.0959	0.0739	0.0190
QRNN-U-MIDAS	**0.2004**	**0.0448**	**0.0224**	**0.0050**
ANN-U-MIDAS	0.3754	0.0613	0.0433	0.0106

NB: The first column presents the competing models while the second row presents the metric for 90% quantiles.

dence interval. Figure 5 shows both the observed and in-sample predictions of GDP for the period of 10 years.

For computing the root mean square error for the quantile (RMSEτ) and the mean absolute percentage error of the quantile (MAPEτ), see equations (20) and (21), in the spirit of Zang *et al.* [63], where $y_t(\tau)$ represents the projected values and ρτ() represent the check-function. The DM test is also used to compare the model's correctness. The period is from 1994Q1 to 2022Q1 with a size of 113 quarters. Three alternative high-frequency forecast horizons (hi = 1/12, 12/12, and

Table 6. Variables and their descriptions.

Variables Abbreviation	Description	Frequency
GDP	Gross domestic products	Quarterly
IPI	Industrial price index	monthly
CPI	Consumer price index	monthly
M2	Real M2 money stock	Monthly
PPI	Producer price index	monthly
House	National house price index	Monthly
Unemp	Unemployment . . .	Monthly
S&P500	S&P common stock index	weekly

Table 7. Variables and their descriptions.

Variables Abbreviation	Obs.	mean	min	max	description	Frequency
GDP	113	69556.3	37756	106271.2	Gross domestic products	Quarterly
IPI	340	100.057	94.68	108.38	Industrial price index	monthly
CPI	341	98.69515	95.91	103.44	Consumer price index	monthly
M2	500	4030.701	2305.1	7682.4	Real M2 money stock	Monthly
House	500	100.21	94.43	111.12	National house price index	Monthly
Unemp	500	4.067	2.30	5.80	Unemployment	Monthly
S&P500	1400	7599.63	3437.4	14289.4	S&P common stock index	weekly

Source: FRED (Federal Reserve Economic Database, https://fred.stlouisfed.org/).

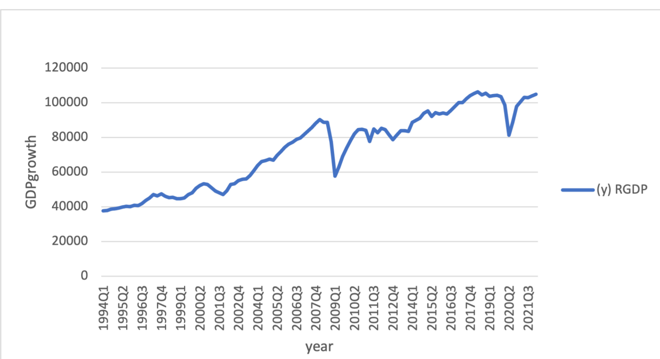

Figure 3. Quarterly GDP growth rate: Presents the time series of GDP.

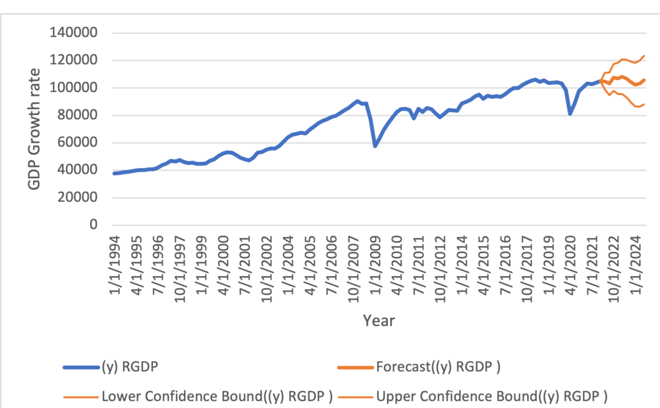

Figure 4. Forcast of quarterly GDP: presents three years or twelve quarters forecast of GDP.

24/12) are used to generate the RMSEs and MAPEs for each rolling, which relate to low-frequency forecast horizons (h = 0,

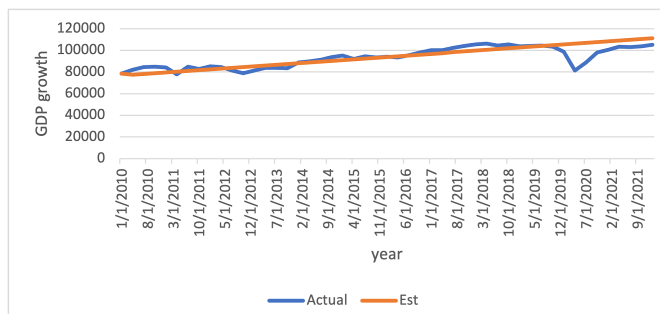

Figure 5. In-sample forcast of GDP growth: presents an in-sample forecast for the period of ten (10) year.

1, and 2). In this application, we identify five sample quantiles that are considered: 0.1, 0.25, 0.5, 0.75, and 0.9.

Table 8 lists the GACV's recommended settings for the optimal lag order and the required number of inner nodes. It is important to notice that the chosen best values in this QRNN-U-MIDAS model are essentially close as those in the ANN-U-MIDAS model and that the optimal parameters are quite similar across different quantiles. Test output from seasonally adjusted variables presented in Table 4 are presented below, and it has shown favourable results in respect of our novel model across all quantiles.

Tables 9–13 report the RMSE and MAPE results of the U-MIDAS, MIDAS, ANN-U-MIDAS, and QRNN-U-MIDAS models along five quantiles (τ) of 0.10, 0.25, 0.50, 0.75, and 0.90. The first column represents the forecasting horizons (h_i). The columns of U-MIDAS, MIDAS, ANN-U-MIDAS, and QRNN-U-MIDAS report the average of RMSE and MAPE across rolling windows. We report the DM test across all mod-

Table 8. Best values of lags and number of nodes in the inner layer.

	U-MIDAS	GACV	MIDAS	GACV	ANN-U-MIDAS	GACV	QRNN-U-MIDAS	GACV
τ/L	l_1, l_2, l_3, l_4		l_1, l_2, l_3, l_4		l_1, l_2, l_3, l_4,**J**		l_1, l_2, l_3, l_4,**J**	
0.10	2, 3, 4, 9	0.0021	3, 3, 6, 9	0.0018	6, 6, 9, 9,10	0.0020	5, 6, 9, 9,10	0.0020
0.25	2, 3, 4, 9	0.0017	3, 3, 6, 9	0.0017	6, 6, 9, 9,10	0.0016	5, 6, 9, 9,10	0.0011
0.50	2, 6, 5, 9	0.0013	3, 3, 5, 9	0.0012	5, 5, 9, 9,10	0.0016	5, 6, 9, 9,10	0.0015
0.75	3, 6, 5, 8	0.0018	3, 3, 5, 9	0.0015	6, 6, 9, 8,10	0.0018	5, 6, 9, 9,10	0.0018
0.90	3, 6, 5, 8	0.0023	3, 3, 6, 9	0.0018	6, 6, 9, 9,10	0.0019	5, 6, 9, 9,10	0.0007

NB: number of lag orders for the four variables used in the model and inner layer parameter.

Table 9. RMSE and MAPE for Japan's quarterly GDP along three forecast horizons $\tau = 0.1$.

	Training data Performance measures at 0.1									
	RMSE					MAPE				
hi	U-MIDAS	MIDAs	ANN-U-MIDAS	QRNN-U-MIDAS	DM test	U-MIDAS	MIDAS	ANN-U–MIDAS	QRNN-U-MIDA	DM test
1/3	0.6935	0.9488	5.8031	**0.212**	5.243 (**0.000**)	0.1372	0.1824	1.1502	**0.0177**	4.264 (**0.000**)
2/3	0.5842	1.0013	4.8231	**0.5562**	7.140 (**0.000**)	0.0912	0.2435	0.3689	**0.0094**	8.022 (**0.000**)
3/3	0.6007	0.7361	4.7612	**0.4356**	3.025 (**0.012**)	0.1701	0.2006	1.0378	**0.1029**	5.386 **(0.000)**
	Testing data Performance measures at 0.1									
	RMSE					MAPE				
hi	U-MIDAS	MIDAS	ANN-U-MIDA	QRNN-U-MIDA	DM test	U-MIDAS	MIDAS	ANN-U-MIDAS	QRNN-U-MIDA	DM test
1/3	0.5167	0.8377	4.0291	**0.183**	5.382 (**0.000**)	0.1291	0.1965	1.0351	**0.011**	4.295 (**0.000**)
2/3	0.4651	0.9103	3.8231	**0.0109**	3.217 (**0.012**)	0.0992	0.2001	0.7156	**0.0092**	4.237 (**0.000**)
3/3	0.6103	0.5361	5.1512	**0.2011**	10.123 (**0.000**)	0.1808	**0.1722**	1.2072	0.1756	7.001 (**0.000**)

NB. The first column represents the forecasting horizons (h_i). The average of RMSE and MAPE for all metrics with the respective DM test across all models is reported, with the p-value in parenthesis and in bold.

els, for example, U-MIDAS model 1 as e_i^2 and QRNN-U-MIDAS model 2 as $e'_i{}^2$, in the last column, which includes the DM values together with the p-value in parenthesis and bold; the best outcome is indicated when comparing using boldface. We can infer the following information from these tables: First, all four models had improved RMSE and MAPE on both training and test data sets.

Second, our QRNN-U-MIDAS model beats the U-MIDAS, MIDAS, and ANN-U-MIDAS models on test data and training data. The former has lowest RMSE and MAPE average values. Third, given that the DM test findings are statistically significant at the 5% level, QRNN-U-MIDAS has proven superior in this case. As a result, the proposed QRNN-U-MIDAS model works well across all forecast horizons and representative quantiles and achieves reasonably high prediction accuracy. We also confirmed the workability of the developed "QRNN-UMIDASV3" package on the set of real-world data for both Japan and the United States. However, results from the US data have not been reported because of space.

5. Conclusion

In order to investigate the heterogeneous nonlinear interaction between variables on mixed sample frequency data, we develop a unique hybrid QRNN-U-MIDAS model. The U-MIDAS approach is incorporated into the QRNN framework to create the QRNN-U-MIDAS model. The proposed QRNN-U-MIDAS model was implemented in an R-package that we created to perform both simulation and empirical data applications. The numerical findings from simulation and empirical data demonstrate that the QRNN-U-MIDAS model is capable of exploring the nonlinear interaction between variables on mixed data sampling frequency. Additionally, in terms of accuracy and predictability, it greatly outperforms a number of competing models.

The following key benefits of the suggested QRNN-U-MIDAS model are; utilizing all of the information present in the original data enables us to directly model on mixed sample frequency data and aids in enhancing forecast accuracy. Second,

Table 10. RMSE and MAPE for Japan's quarterly GDP along three forecast horizons $\tau = 0.25$.

	Training data Performance measures at 0.25									
	RMSE					MAPE				
hi	U-MIDA	MIDAs	ANN-U-MIDAS	QRNN-U-MIDA	DM test	U-MIDAS	MIDAS	ANN-U–MIDAS	QRNN-U-MIDA	DM test
1/3	4.8096	9.0024	3.1578	**2.0843**	5.108 (**0.000**)	0.5168	0.7395	4.7614	**0.024**	6.101 (**0.000**)
2/3	1.6923	1.5188	5.8231	**0.5562**	10.135 (**0.000**)	0.1824	0.2297	0.3689	**0.0136**	5.126 (**0.000**)
3/3	0.5168	0.7391	4.9012	**0.4356**	2.839 (**0.030**)	0.1971	0.312	1.0378	**0.0929**	2.920 **(0.026)**
	Testing data Performance measures at 0.25									
	RMSE					MAPE				
hi	U-MIDAS	MIDAs	ANN-U-MIDA	QRNN-U-MIDA	DM test	U-MIDAS	MIDAS	ANN-U-MIDAS	QRNN-U-MIDA	DM test
1/3	4.532	7.2201	2.3356	**1.7521**	3.212 (**0.012**)	0.1092	0.1865	1.1451	**0.017**	2.846 (**0.027**)
2/3	1.8601	1.0133	4.8235	**0.511**	7.478 **(0.000)**	0.2101	0.2321	0.1154	**0.082**	5.116 **(0.000)**
3/3	0.8214	0.9573	5.1411	**0.2301**	8.317 (**0.000**)	0.1608	0.1823	**0.0072**	0.1756	6.234 (**0.000**)

NB: See details to table in Table 7.

Table 11. RMSE and MAPE for Japan's quarterly GDP along three forecast horizons $\tau = 0.5$.

	Training data Performance measures at 0.5									
	RMSE					MAPE				
hi	U-MIDA	MIDAs	ANN-U-MIDAS	QRNN-U-MIDA	DM test	U-MIDAS	MIDAS	ANN-U–MIDAS	QRNN-U-MIDA	DM test
1/3	0.6935	0.9202	5.8578	**0.5301**	2.375 (**0.032**)	0.3008	0.5375	4.7114	**0.0201**	5.387 (**0.000**)
2/3	0.5168	0.7392	4.9123	**0.2371**	5.426 (**0.001**)	0.1921	0.2208	0.4119	**0.1891**	3.265 (**0.012**)
3/3	0.1271	0.1824	1.2425	**0.0856**	6.035 (**0.000**)	0.1054	0.232	1.0111	**0.0907**	10.005 **(0.000)**
	Testing data Performance measures at 0.5									
	RMSE					MAPE				
hi	U-MIDAS	MIDAs	ANN-U-MIDA	QRNN-U-MIDA	DM test	U-MIDAS	MIDAS	ANN-U-MIDAS	QRNN-U-MIDA	DM test
1/3	0.5998	0.7271	4.3216	**0.5080**	8.576 (**0.000**)	0.2032	0.3165	1.0325	**0.0197**	5.027 (**0.000**)
2/3	0.6174	0.8123	5.0245	**0.3132**	3.254 **(0.012)**	0.1801	0.272	0.196	**0.0082**	2.903 **(0.032)**
3/3	0.2017	**0.1573**	1.1901	0.1701	7.387 (**0.000**)	0.1298	0.1903	0.0912	**0.0746**	10.839 (**0.001**)

NB: See details to table in Table 7.

the QRNN-MIDAS model, which leverages the power of quantile regression, effectively captures the full conditional distribution of the estimate of a response variable. Third, by incorporating the potent capability of neural networks in solving nonlinear issues, the QRNN-U-MIDAS model may be employed to investigate the complex nonlinear pattern. From the findings, the accuracy measured across the conditionally distributed quantiles of the preliminary results indicates that QRNN-U-MIDAS presents the minimum error over the competing models in most forecast horizons.

Table 12. RMSE and MAPE for Japan's quarterly GDP along three forecast horizons τ = 0.75.

	Training data Performance measures at 0.75									
	RMSE					MAPE				
hi	U-MIDA	MIDAs	ANN-U-MIDAS	QRNN-U-MIDA	DM test	U-MIDAS	MIDAS	ANN-U–MIDAS	QRNN-U-MIDA	DM test
1/3	0.6023	0.7901	5.5121	**0.4604**	3.108 (**0.012**)	0.3834	0.5756	5.1212	**0.0190**	3.278 (**0.013**)
2/3	0.4868	0.6835	4.729	**0.2831**	2.106 (**0.034**)	0.2102	0.2319	0.5631	**0.0919**	2.210 (**0.004**)
3/3	0.2151	0.2994	0.9915	**0.0783**	1.439 (**0.006**)	0.1078	0.2606	1.1014	**0.0735**	3.286 2(**0.013**)
	Testing data Performance measures at 0.75									
	RMSE					MAPE				
hi	U-MIDAS	MIDAs	ANN-U-MIDA	QRNN-U-MIDA	DM test	U-MIDAS	MIDAS	ANN-U-MIDAS	QRNN-U-MIDA	DM test
1/3	0.5781	0.7362	4.2303	**0.508**	6.325 (**0.000**)	0.2231	0.3143	1.0325	**0.0197**	8.132 (**0.000**)
2/3	0.5011	0.7324	4.922	**0.2017**	3.326 (**0.012**)	0.1801	0.28	0.2157	**0.0942**	5.146 (**0.000**)
3/3	0.2383	0.2045	1.1901	**0.1802**	5.213 (**0.000**)	0.1239	0.1826	0.1863	**0.1011**	2.921 (**0.022**)

NB: See details to table in Table 7.

Table 13. RMSE and MAPE for Japan's quarterly GDP along three forecast horizons τ= 0.9.

	Training data Performance measures at 0.9									
	RMSE					MAPE				
hi	U-MIDA	MIDAs	ANN-U-MIDAS	QRNN-U-MIDA	DM test	U-MIDAS	MIDAS	ANN-U–MIDAS	QRNN-U-MIDA	DM test
1/3	0.6131	0.8629	4.8174	**0.4906**	4.576 (**0.000**)	0.3101	0.4132	4.1012	**0.0193**	5.852 (**0.000**)
2/3	0.5002	0.7183	3.925	**0.2582**	2.154 (**0.000**)	0.1721	0.1397	0.5328	**0.0917**	5.249 (**0.000**)
3/3	0.1257	0.1114	1.0293	**0.0918**	10.839 (**0.000**)	0.0901	0.1243	0.9492	**0.0838**	2.005 (**0.0034**)
	Testing data Performance measures at 0.9									
	RMSE					MAPE				
hi	U-MIDAS	MIDAs	ANN-U-MIDA	QRNN-U-MIDA	DM test	U-MIDAS	MIDAS	ANN-U-MIDAS	QRNN-U-MIDA	DM test
1/3	0.5017	0.7813	4.3191	**0.5144**	5.279 (**0.000**)	0.3102	0.3502	1.1305	**0.0163**	7.108 (**0.000**)
2/3	0.4283	0.8032	5.128	**0.2146**	7.042 (**0.000**)	0.2001	0.252	0.1861	**0.0899**	3.285 (**0.012**)
3/3	0.2018	0.1369	1.1654	**0.0791**	2.133 (**0.031**)	0.0961	0.1307	**0.0975**	0.0993	10.132 (**0.000**)

NB: See details to table in Table 7.

The proposed model, "QRNN-U-MIDAS," is a novel model that will give decision-makers a precise estimate of macroeconomic variable, like GDP. Moreover, the developed package "QRNNUMIDASV3" shall be sent to CRAN for other researchers to use and improve on it.

Acknowledgment

The first author would like to first thank Allah the Almighty for keeping him alive and healthy to undertake this research. He wishes to acknowledge with appreciation the prompt response and academically sound guidance he enjoyed from his supervisor, Associate Professor Mohd Tahir Ibn Ismail, at the School of Mathematics, Universiti Sains Malaysia. He also wish to appreciate the Tertiary Education Trust Fund (TETFund) for sponsoring not only this paper but his entire study at the Universiti Sains Malaysia (Award No. TETF/ES/POLY/YOBE STATE/TSAS/2019). He thanks the Department of Statistics, Federal Polytechnic Damaturu for granting him this opportunity.

References

[1] D. T. Utari & H. Ilma, "Comparison of methods for mixed data sampling (MIDAS) regression models to forecast Indonesian GDP using agricultural exports", AIP Conference Proceedings **2021** (2018) 060016. https://doi.org/10.1063/1.5062780

[2] E. Andreou, E. Ghysels & A. Kourtellos, "Regression models with mixed sampling frequencies", Journal Econometrics **158** (2010) 246. https://doi.org/10.1016/j.jeconom.2010.01.004

[3] E. Ghysels, A. Sinko & R. Valkanov, "MIDAS regressions: Further results and new directions", Econometric Reviews **26** (2007) 53. https://doi.org/10.1080/07474930600972467

[4] E. Ghysels, P. Santa-Clara & R. Valkanov, "The MIDAS Touch: Mixed data sampling regression models", CIRANO Working Papers (2004). https://rady.ucsd.edu/_files/faculty-research/valkanov/midas-touch.pdf

[5] C. Foroni & C. Schumacher, *U-MIDAS : MIDAS regressions with unrestricted lag polynomials Massimiliano Marcellino Discussion Paper Series 1 : Economic Studies*, Deutsche Bundesbank **35** (2011). https://hdl.handle.net/1814/40284

[6] D. Pradeepkumar & V. Ravi, "Forecasting financial time series volatility using Particle Swarm Optimization trained Quantile Regression Neural Network", Applied Soft Computing Journal **58** (2017) 35. https://doi.org/10.1016/j.asoc.2017.04.014

[7] C. Foroni, M. Marcellino & C. Schumacher, *Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials*, (2014), [Online]. Available: http://wileyonlinelibrary.com/journal/rss-datasets

[8] E. Ghysels, V. Kvedaras & V. Zemlys, "Mixed frequency data sampling regression models: The R package midasr", J Stat Softw. **72** (2016) 1. https://doi.org/10.18637/jss.v072.i04

[9] F. Barsoum & S. Stankiewicz, "Forecasting GDP growth using mixed-frequency models with switching regimes", Int J Forecast. **31** (2015) 33. https://doi.org/10.1016/j.ijforecast.2014.04.002

[10] Q. Xu, S. Liu, C. Jiang & X. Zhuo, "QRNN-MIDAS: A novel quantile regression neural network for mixed sampling frequency data", Neurocomputing **457** (2021) 105. https://doi.org/10.1016/j.neucom.2021.06.006

[11] Q. Xu, X. Zhuo, C. Jiang & Y. Liu, "An artificial neural network for mixed frequency data," Expert Syst Appl. **118** (2019) 139. https://doi.org/10.1016/j.eswa.2018.10.013

[12] R. K. B. Navas, S. Prakash & T. Sasipraba, "Artificial neural network based computing model for wind speed prediction: A case study of Coimbatore, Tamil Nadu, India", Physica A: Statistical Mechanics and its Applications **542** (2020) 123383. https://doi.org/10.1016/j.physa.2019.123383

[13] L. Breiman, "Heuristics of instability and stabilization in model selection", Ann. Statist. **24** (1996) 2350. https://doi.org/10.1214/aos/1032181158

[14] Y. He & W. Zhang, "Probability density forecasting of wind power based on multi-core parallel quantile regression neural network", Knowl Based Syst. **209** (2020) 1. https://doi.org/10.1016/j.knosys.2020.106431

[15] Š. Lyócsa & D. Stašek, "Improving stock market volatility forecasts with complete subset linear and quantile HAR models", Expert Syst Appl. **183** (2021) 115416. https://doi.org/10.1016/j.eswa.2021.115416

[16] P. H. Franses, "Yet another look at MIDAS regression", Econometric Institute Research Papers (2016) 1. https://repub.eur.nl/pub/93331/EI2016-32.pdf

[17] M. T. Armesto, K. M. Engemann & M. T. Owyang, "Forecasting with mixed frequencies", Federal Reserve Bank of St. Louis Review **92** (2010) 536. https://doi.org/10.20955/r.92.521-36

[18] Q. Xu, L. Wang, C. Jiang & X. Zhang, "A novel UMIDAS-SVQR model with mixed frequency investor sentiment for predicting stock market volatility", **132** (2019) 12. https://doi.org/10.1016/j.eswa.2019.04.066

[19] Y. Liu, *Statistical methods for mixed frequency data sampling models*, Digital Commons @ Michigan Tech, Michigan Technological University (2019). https://doi.org/10.37099/mtu.dc.etdr/881

[20] M. Damane, "Forecasting the Government of Lesotho's budget: an AR-MIDAS approach", African J. of Economic and Sustainable Development **7** (2020) 256. https://doi.org/10.1504/ajesd.2020.105688

[21] D. Bams, G. Blanchard, I. Honarvar & T. Lehnert, "Does oil and gold price uncertainty matter for the stock market?", J Empir Finance. **44** (2017) 285. https://doi.org/10.1016/j.jempfin.2017.07.003

[22] C. Baumeister, P. Guérin & L. Kilian, "Do high-frequency financial data help forecast oil prices? The MIDAS touch at work", Int J Forecast. **31** (2015) 238. https://doi.org/10.1016/j.ijforecast.2014.06.005

[23] S. Gunay, G. Can & M. Ocak, "Forecast of China's economic growth during the COVID-19 pandemic: a MIDAS regression analysis", Journal of Chinese Economic and Foreign Trade Studies **14** (2021) 3. https://doi.org/10.1108/JCEFTS-08-2020-0053

[24] T. B. Götz and K. Hauzenberger, "Large mixed-frequency VARs with a parsimonious time-varying parameter structure," Econom J. **24** (2021) 442. https://doi.org/10.1093/ectj/utab001

[25] S. Das, R. Demirer, R. Gupta & S. Mangisa, "Ac ce pt cr t", Structural Change and Economic Dynamics (2019). https://doi.org/10.1016/j.strueco.2019.05.007

[26] A. Babii, E. Ghysels & J. Striaukas, "Machine learning time series regressions with an application to nowcasting", Journal of Business and Economic Statistics **40** (2022) 1094. https://doi.org/10.1080/07350015.2021.1899933

[27] J. Khoo & A. W. K. Cheung, "Does geopolitical uncertainty affect corporate financing? Evidence from MIDAS regression", Global Finance Journal **47** (2021) 100519. https://doi.org/10.1016/j.gfj.2020.100519

[28] S. Penev, D. Leonte, Z. Lazarov & R. A. Mann, "Applications of MIDAS regression in analysing trends in water quality", J Hydrol (Amst). **511** (2014) 159. https://doi.org/10.1016/j.jhydrol.2014.01.031

[29] X. Zhao, M. Han, L. Ding & A. C. Calin, "Forecasting carbon dioxide emissions based on a hybrid of mixed data sampling regression model and back propagation neural network in the USA", Environmental Science and Pollution Research **25** (2018) 2899. https://doi.org/10.1007/s11356-017-0642-6

[30] S. Bhaghoe, G. Ooft & P. H. Franses, "Estimates of quarterly GDP growth using MIDAS regressions", Econometric Institute Research Papers EI2019-29, Erasmus University Rotterdam, Erasmus School of Economics (ESE), Econometric Institute (2019). https://repub.eur.nl/pub/118667/EI2019-29-Report.pdf

[31] Z. Pan, Q. Wang, Y. Wang & L. Yang, "Forecasting U.S. real GDP using oil prices: A time-varying parameter MIDAS model", Energy Econ. **72** (2018) 177. https://doi.org/10.1016/j.eneco.2018.04.008

[32] H. Hassani, A. Rua, E. S. Silva & D. Thomakos, "Monthly forecasting of GDP with mixed-frequency multivariate singular spectrum analysis", Int J Forecast. **35** (2019) 1263. https://doi.org/10.1016/j.ijforecast.2019.03.021

[33] N. Kingnetr, T. Tungtrakul & S. Sriboonchitta, *Forecasting GDP growth in Thailand with different leading indicators using MIDAS regression models*, Part of the Studies in Computational Intelligence book series (SCI, volume 692) (2017). https://doi.org/10.1007/978-3-319-50742-2

[34] G. L. Mazzi & J. Mitchell, "New methods for timely estimates: nowcasting euro area GDP growth using quantile regression", Statistical Working Paper (2020). https://doi.org/10.2785/26603

[35] R. S. Mariano & S. Ozmucur, "Predictive performance of mixed-frequency nowcasting and forecasting models (with application to Philippines inflation and GDP growth)", Journal of Quantitative Economics **19** (2021) 383. https://doi.org/10.1007/s40953-021-00276-6

[36] S. A. J. Rani, V. V. Haragopal & M. K. Reddy, "Forecasting inflation rate of India using neural networks", International Journal of Computer Applications **158** (2017) 45. https://www.ijcaonline.org/archives/volume158/number5/rani-2017-ijca-912866.pdf

[37] G. Kapetanios, "Analysis of the most recent modelling techniques for big data with particular attention to Bayesian ones", (2018). https://doi.org/10.2785/679114

[38] A. F. Adedotun, T. Latunde & O. A. Odusanya, "Modelling and forecast-

ing climate time series with state-space model", J. Nig. Soc. Phys. Sci. **2** (2020) 94. https://doi.org/10.46481/jnsps.2020.94
[39] G. O. Odekina, A. F. Adedotun, and O. F. Imaga, "Modeling and forecasting the third wave of Covid-19 incidence rate in Nigeria using vector autoregressive model approach", J. Nig. Soc. Phys. Sci. **4** (2022) 431. https://doi.org/10.46481/jnsps.2021.431
[40] E. M. Chouit, M. Rachdi, M. Bellafkih & B. Raouyane, "Forecasting of the epidemiological situation: Case of COVID-19 in Morocco", J. Nig. Soc. Phys. Sci. textbf4 (2022) 843. https://doi.org/10.46481/jnsps.2022.843
[41] T. H. Le, "Forecasting value at risk and expected shortfall with mixed", Int J Forecast. **36** (2022) 1362. https://doi.org/10.1016/j.ijforecast.2020.01.008
[42] Y. Wei, Q. Yu, J. Liu & Y. Cao, "Hot money and China's stock market volatility: Further evidence using the GARCH–MIDAS model", Physica A: Statistical Mechanics and its Applications **492** (2018) 923. https://doi.org/10.1016/j.physa.2017.11.022
[43] A. Amendola, V. Candila & G. M. Gallo, "Choosing the frequency of volatility components within the Double Asymmetric GARCH–MIDAS–X model", Econom Stat. **20** (2021) 12. https://doi.org/10.1016/j.ecosta.2020.11.001
[44] L. Renato, F. Meng & L. Godeiro, "Quantile forecasting with mixed-frequency data", Int J Forecast. **36** (2020) 1149. https://doi.org/10.1016/j.ijforecast.2018.09.011
[45] E. Ghysels & H. Qian, "Estimating MIDAS regressions via OLS with polynomial parameter profiling", Econom Stat. **9** (2019) 1. https://doi.org/10.1016/j.ecosta.2018.02.001
[46] E. Ghysels, V. Kvedaras & V. Zemlys-Balevičius, "Mixed data sampling (MIDAS) regression models", Handbook of Statistics **42** (2020) 117. https://doi.org/10.1016/bs.host.2019.01.005
[47] E. Ghysels, V. Kvedaras & V. Zemlys, "Mixed frequency data Sampling regression models: The R Package midasr", Journal of Statistical Software **72** (2016) 14. https://doi.org/10.18637/jss.v072.i04
[48] G. Kapetanios & F. Papailias, "Big data & macroeconomic nowcasting: Methodological review", Economic Statistics Centre of Excellence, National Institute of Economic and Social Research (2018). https://escoe-website.s3.amazonaws.com/wp-content/uploads/2020/07/13161005/ESCoE-DP-2018-12.pdf
[49] Y. Zhang, C. L. Yu & H. Li, "Nowcasting GDP using dynamic factor model with unknown number of factors and stochastic volatility: A Bayesian approach", Econom Stat. **24** (2022) 75. https://doi.org/10.1016/j.ecosta.2021.08.009
[50] C. Foroni, M. Marcellino & D. Stevanovic, "Forecasting the Covid-19 recession and recovery: Lessons from the financial crisis", Int J Forecast. **38** (2021) 596. https://doi.org/10.1016/j.ijforecast.2020.12.005
[51] A. Richardson, T. V. F. Mulder & T. Vehbi, "Nowcasting New Zealand GDP using machine learning algorithms", IFC Bulletins chapters, in: Bank for International Settlements (ed.), The use of big data analytics and artificial intelligence in central banking, Bank for International Settlements **50** (2018) 23. https://www.bis.org/ifc/publ/ifcb50_15.pdf
[52] V. Umarani, A. Julian & J. Deepa, "Sentiment analysis using various machine learning and deep learning techniques", J. Nig. Soc. Phys. Sci. **3** (2021) 308. https://doi.org/10.46481/jnsps.2021.308
[53] G. Ramadani, M. Petrovska & V. Bucevska, "Evaluation of mixed frequency approaches for tracking near-term economic developments in North Macedonia", South East European Journal of Economics and Business **16** (2021) 43. https://doi.org/10.2478/jeb-2021-0013
[54] S. M. Eds & R. Goebel, *Multiple-Aspect of Semantic Trajectories*, First International Workshop, MASTER 2019, Held in Conjunction with ECML-PKDD, Würzburg, Germany Proceedings (2019). https://doi.org/10.1007/978-3-030-38081-6
[55] C. Challu, K. G. Olivares, G. Welter & A. Dubrawski, "DMIDAS: Deep mixed data sampling regression for long multi-horizon time series forecasting", (2021). https://doi.org/10.48550/arXiv.2106.05860
[56] K. Benidis *et al.*, "Neural forecasting : Introduction and literature overview", (2020). https://doi.org/10.1145/3533382
[57] R. K. B. Navas, S. Prakash, and T. Sasipraba, "Artificial neural network based computing model for wind speed prediction: A case study of Coimbatore, Tamil Nadu, India", Physica A: Statistical Mechanics and its Applications **542** (2020) 123383. https://doi.org/10.1016/j.physa.2019.123383
[58] H. Mo, J. Wang & H. Niu, "Exponent back propagation neural network forecasting for financial cross-correlation relationship", Expert Syst Appl. **53** (2016) 106. https://doi.org/10.1016/j.eswa.2015.12.045
[59] Q. Xu, K. Deng, C. Jiang, F. Sun & X. Huang, "Composite quantile regression neural network with applications", Expert Syst Appl. **76** (2017) 129. https://doi.org/10.1016/j.eswa.2017.01.054
[60] A. J. Cannon, "Neural networks for probabilistic environmental prediction: Conditional Density Estimation Network Creation and Evaluation (CaDENCE) in R", Comput Geosci. **41** (2012) 126. https://doi.org/10.1016/j.cageo.2011.08.023
[61] S. Galeshchuk, "Neural networks performance in exchange rate prediction", Neurocomputing **172** (2016) 446. https://doi.org/10.1016/j.neucom.2015.03.100
[62] M. Stevanović, S. Vujičić, and A. M. Gajić, "Gross domestic product estimation based on electricity utilization by artificial neural network", Physica A: Statistical Mechanics and its Applications **489** (2018) 28. https://doi.org/10.1016/j.physa.2017.07.023
[63] W. Zhang, H. Quan, and D. Srinivasan, "An improved quantile regression neural network for probabilistic load forecasting", IEEE Trans Smart Grid **10** (2019) 4425. https://doi.org/10.1109/TSG.2018.2859749
[64] S. Asimakopoulos, J. Paredes, and T. Warmedinger, "Forecasting fiscal time series using mixed frequency data", ECB Working Paper **1550** (2013) 1. https://www.ecb.europa.eu//pub/pdf/scpwps/ecbwp1550.pdf
[65] T. Alam, "Forecasting exports and imports through artificial neural network and autoregressive integrated moving average", Decision Science Letters **8** (2019) 249. https://doi.org/10.5267/j.dsl.2019.2.001
[66] F. X. Diebold & R. S. Mariano, "Comparing predictive accuracy", Journal of Business and Economic Statistics **20** (2002) 134. https://doi.org/10.1198/073500102753410444