From MIDAS to Deep Learning: A comprehensive benchmark of big data economic forecasting models

Miguel Benítez Humanes

Universitat Politècnica de Catalunya

January 23, 2024

These slides, the main text, the data and the codes used in this thesis are all publicly available in the Github repository:
`https://github.com/miguelBhumanes/FinalMasterThesis.git`

In this thesis...

- Initial version of a comprehensive benchmark of economic forecasting models
- Testing a models and refinements of these models to achieve a higher degree of comprehensiveness
- Facilitate research journey of economic forecasters

Agenda

1. Background and Motivation
2. Forecasting Models Assessed
3. Data
4. Results
5. Insights and Limitations

The need for economic forecasting

2019 **2023**

3.00 € 9.90 €

Source: *(Images generated with GPT4)*

Literature on Economic Forecasting

- Extense (Ranging from mostly economic papers to pure Mathematics and Computer Science)
- Many different models. Preference for linear models.
- **Typical paper structure:**
  - Introduces slight methodological innovation
  - Or uses different inputs (usually adding a specific indicator)
  - Computes the RMSE of the model and its immediate simpler version
- **Problems:**
  - The data sets used are rarely standard
  - The methods for training very similar models vary
  - Methodological improvements found are typically applicable to many existing model architectures, but seldom tried on more than one $\rightarrow$ Numerous research gaps

Highlighted articles (I)

- **Hopp. (2022).**[6]
  - **Pros**
    - Best available benchmark (LR, **MIDAS**, BVAR, DFM, **ANN**, **LSTM**, Tree Methods)
  - **Cons**
    - Focuses on nowcasting
    - Does not consider inflation
    - Uses only economic variables
- **Bok et al. (2018).**[2]
  - **Pros**
    - Relatively simple model Dynamic Factor Model (DFM) [9] with good results
  - **Cons**
    - Does not explore variations of the DFM proposed

Highlighted articles (II)

- **Ardia et al. (2019).**[6]
  - Rich set of data, including financial and text variables
  - Regularization to determine best inputs
  - Use of MIDAS

Research Gap → Combining the benchmark approach of Hopp. (2022) with the in-depth assessment of models in Ardia et al. (2019).

Objective of the thesis

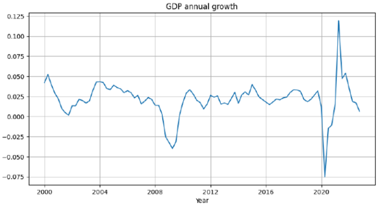
Research Objective: creating sufficiently comprehensive benchmark. Useful as initial source in economic forecasting research:

1. Assessing both GDP growth and inflation forecasting ability
2. Considering different horizons and economic contexts
3. Including many different model architectures
4. Using large numbers of predictors (including text based indicators)
5. Considering MIDAS methods to deal with frequency alignment

Model "families" in the benchmark

1. Autoregressive Models (ARs)
2. Dynamic Factor Models (DFMs)
3. Neural Network Models (NNs)

AR Models: Stationarity

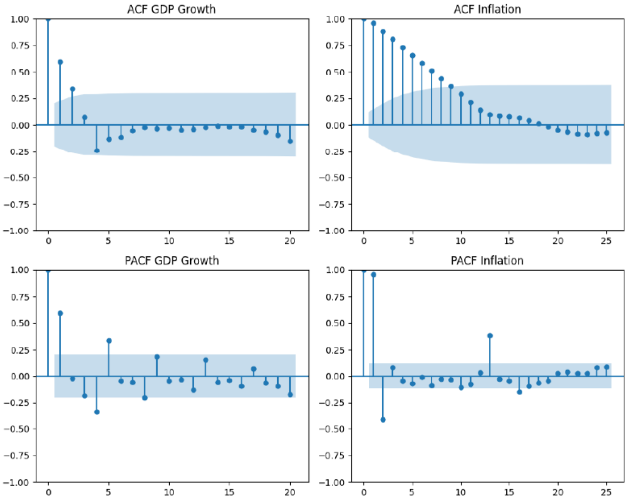


	ADF Test	KPSS Test
GDP Growth	0.1769	0.1000*
Inflation	0.1789	0.1000*
GDP Growth (ex. 2020-2022)	0.1349	
Inflation (ex. 2020-2022)	0.0418	

Table 1: ADF and KPSS p-values for GDP Growth and Inflation

*Note: KPSS p-values are higher than 0.1 but the Python package used (statsmodels) does not display the exact result above that threshold.

AR Models: Model identification



AR Models: All models and results

	AR	Auto ARIMA	LSTM
GDP Growth	4.00 AR(5)	3.55 AR(1)	3.53
GDP Growth (ex. Covid)	0.36 AR(5)	0.36 AR(5)	0.46
Inflation	1.14 AR(13)	1.21 ARMA(1,3)	2.64
Inflation (ex. Covid)	0.50 AR(13)	0.41 AR(3)	0.47

Table 2: One Quarter Ahead RMSE (%)

	AR	Auto ARIMA	LSTM
GDP Growth	5.24 AR(5)	5.09 AR(1)	3.73
GDP Growth (ex. Covid)	0.85 AR(5)	0.85 AR(5)	0.64
Inflation	3.81 AR(13)	3.29 ARMA(1,3)	3.30
Inflation (ex. Covid)	0.73 AR(13)	0.42 AR(3)	0.45

Table 3: One Year Ahead RMSE (%)

DFM: Introduction

- DFMs and Vector Autoregressions (VARs) are the most popular forecasting models in economics [9].
- DFMs are simpler, while the results are very similar to the ones of Bayesian VARs (BVARs) [3]

$$Y_t = f(F_t, F_{t-1}, F_{t-2}, ...) + V_t \quad , \quad V_t \sim N(0, R)$$

$$F_t = g(F_{t-1}, F_{t-2}, ...) + W_t \quad , \quad W_t \sim N(0, Q)$$

DFM: MIDAS (I)

The factors in that regression will be in different frequencies:

1. **Alternative 1:** Doing the average
2. **Alternative 1:** **Factor MIDAS** (Marcellino and Schumacher, 2010) [7]

DFM: MIDAS (II) [4]

$$y_t = g(\sum_{i=1}^{k} w_i x_{s(t)-i}; \beta) + \epsilon_t = \beta^{'}(\sum_{i=1}^{k} w_i x_{s(t)-i}) + \epsilon_t$$

$$= \beta_0 + \beta_1(\sum_{i=1}^{k} w_i x_{s(t)-i}) + \beta_2(\sum_{i=1}^{k} w_i x_{s(t-1)-i}) + ... + \beta_N(\sum_{i=1}^{k} w_i x_{s(t-N+1)-i}) + \epsilon_t$$

$$\sum_{i=1}^{k} w_i = 1 \quad , \quad w_i = h(i, \theta) \quad , \quad s(t) = \sum_{j=1}^{t} m_j$$

DFM: MIDAS (III)

Equation 3: Normalized Exponential Almon

$$h(i,\theta) = \frac{exp(\theta_1 \cdot i + \theta_2 \cdot i^2)}{\sum_{i=1}^{k} \exp(\theta_1 \cdot i + \theta_2 \cdot i^2)}$$

DFM: Competing models

The models will consider:

1. Different forecasting targets (GDP growth and inflation)
2. Different forecasting horizons (1 quarter and 1 year ahead)
3. Different economic situations (excluding, including the pandemic years)
4. Different sets of variables (economic, financial, text)
5. Two frequency alignment techniques: averages and MIDAS
6. Including, excluding the "blocks" mentioned in Bok et al. (2018)
7. As a substitute for regularization: including only one or the first five PCA factors.

NN: Introduction

- In economics, many relations are non linear [1]
- Neural network architectures are a suitable method to capture them in a scalable way
- **Different types of layers**:
  - Feed Forward (ANN)
  - Long Short Term (LSTM)
  - Gated Recurrent Units (GRU)
  - Self Attention (SA)

NN: Dimensionality reduction

- The neural network could have some variable selection capacity
- In practice, there needs to be variable selection
- Ideally, the selection is done with an autoencoder network [5]
- Here, using the factors for the DFMs

NN: Frequency Alignment

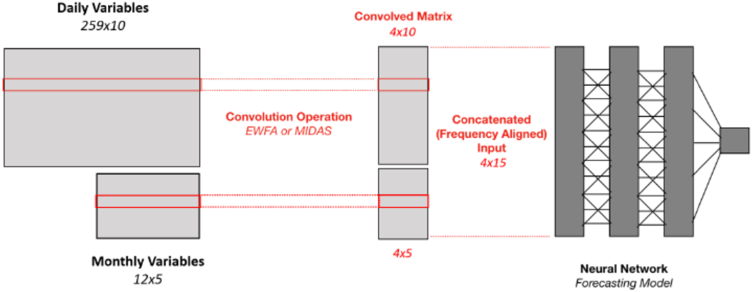


Figure 7: The matrix of daily variables is convolved to obtain a matrix in the monthly frequency. The convolution can use equal weighting (EWFA) or MIDAS weights. The convolved output is then concatenated with the matrix of monthly variables to obtain the frequency aligned input matrix.

NN: Competing models

The models will consider:

1. Different forecasting targets (GDP growth and inflation)
2. Different forecasting horizons (1 quarter and 1 year ahead)
3. The four mentioned types of layers
4. Two frequency alignment techniques: averages and MIDAS
5. Regularization: dropout

Model training

DFMs

- Train-test split 80 - 20
- Rolling window
- MIDAS weights initialized to $\vec{0}$

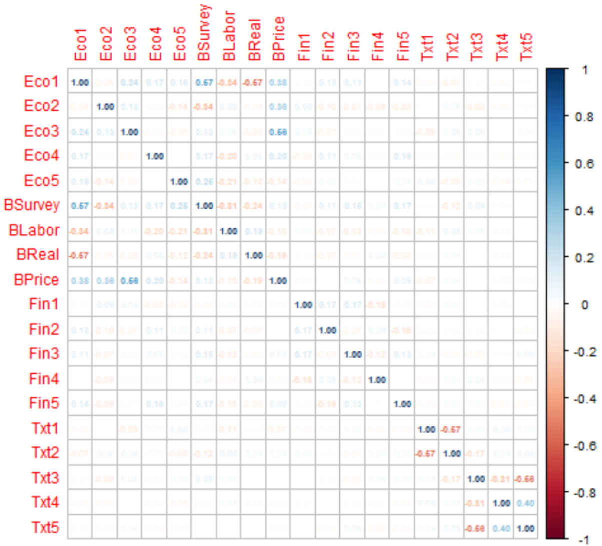
NNs (and AR-LSTM)

- Train-validation-test split 68 - 12 - 20
- Grid search for hyperparameters:
  - Number nodes / layer
  - Number layers
  - Dropout rate
- MIDAS weights initialized to $\vec{0}$, rest randomly
- Adam optimizer (learning rate $= 10^{-4}$)
- Callback to avoid overfitting

Data

- Selected inputs
  - **Economic Variables (44):** slight modification of data set in Bok et al. (2018) [2]
  - **Financial Variables (37):** subset of data set in Stock and Watson (2003) [8]
  - **Text Variables (200):** 200 LDA topics extracted from financial newspapers
- Processing
  - Transformation of all series to stationary form
  - EM-PCA imputation of missing values
  - Five PCA factors for each type of variable
  - Creating of four blocks as indicated in Bok et al. (2018) [2]

Kendall Correlation Matrix



Highlighted Results for GDP Growth

1 Quarter ahead — All sample

	RMSE (%)
ARIMA	3.55
Autoreg. LSTM	3.53
Bok Model	**3.00**
Bok Model (NB)	3.26
E1T1 NB EW	**3.00**
NN GRU MIDAS	3.07

1 Quarter ahead — Ex. Pandemic

	RMSE (%)
ARIMA	**0.36**
Autoreg. LSTM	0.46
Bok Model	0.53
Bok Model (NB)	0.53
E1F1T1 NB EW	0.49

1 Year ahead — All sample

	RMSE (%)
ARIMA	5.09
Autoreg. LSTM	3.73
Bok Model	3.58
Bok Model (NB)	3.85
E5F5 NB EW	**2.90**
E5 NB EW	**2.92**
NN SA EW	3.39

1 Year ahead — Ex. Pandemic

	RMSE (%)
ARIMA	0.85
Autoreg. LSTM	**0.64**
Bok Model	**0.62**
Bok Model (NB)	**0.64**

Highlighted Results for Inflation

1 Quarter ahead

All sample

	RMSE (%)
ARIMA	**1.14**
Autoreg. LSTM	2.64
Bok Model	3.21
Bok Model (NB)	3.14
E5F5 NB MIDAS	2.82
NN ANN MIDAS	2.17

Ex. Pandemic

	RMSE (%)
ARIMA	**0.41**
Autoreg. LSTM	0.47
Bok Model	0.62
Bok Model (NB)	0.64
E1T1 B MIDAS	0.58

1 Year ahead

All sample

	RMSE (%)
ARIMA	3.29
Autoreg. LSTM	3.30
Bok Model	3.12
Bok Model (NB)	3.20
E1F1 B MIDAS	2.91
NN GRU MIDAS	**2.54**

Ex. Pandemic

	RMSE (%)
ARIMA	**0.42**
Autoreg. LSTM	0.45
Bok Model	0.74
Bok Model (NB)	0.64
E1F1 NB MIDAS	0.58
E1T1 NB MIDAS	0.58

Conclusions (I)

Main insights from the benchmark:

- For GDP growth
  1. Bok's model is a very practical forecasting model for GDP [2]
  2. Autoregressive models also obtain good results for GDP
  3. **In volatile economic periods, more complex DFMs might yield better long term forecasts**
- For inflation
  1. ARIMA models appear to be the best forecasters
  2. **In volatile economic periods, appropriate NNs might yield better long term forecasts**

Conclusions (II)

Main insights from the benchmark:

- Additional considerations
  1. In DFMs, text indicators do not appear to improve forecasts consistently
  2. MIDAS alignment appears to worsen forecasts for DFMs for GDP, imrprove them for inflation
  3. For NNs, MIDAS tends to improve forecasts

Conclusions (II)

Main limitations:

- Limited number of included models:
  - Increase hyperparamter space
  - Modifying the data set of inputs
  - Including new architectures (Tree based methods, autoencoder, ...)

- Opportunities to improve regularization methods

References I

Alberto Alesina, Gualtiero Azzalini, Carlo Favero, Francesco Giavazzi, and Armando Miano.
Is it the "how" or the "when" that matters in fiscal adjustments?
IMF Economic Review, 66:144–188, 2018.

Brandyn Bok, Daniele Caratelli, Domenico Giannone, Argia M Sbordone, and Andrea Tambalotti.
Macroeconomic nowcasting and forecasting with big data.
Annual Review of Economics, 10:615–643, 2018.

Christine De Mol, Domenico Giannone, and Lucrezia Reichlin.
Forecasting using a large number of predictors: Is bayesian shrinkage a valid alternative to principal components?
Journal of Econometrics, 146(2):318–328, 2008.

References II

Eric Ghysels, Pedro Santa-Clara, and Rossen Valkanov.
The midas touch: Mixed data sampling regression models.
CIRANO Working Papers, 2004.

Niko Hauzenberger, Florian Huber, and Karin Klieber.
Real-time inflation forecasting using non-linear dimension reduction techniques.
International Journal of Forecasting, 39(2):901–921, 2023.

Daniel Hopp.
Benchmarking econometric and machine learning methodologies in nowcasting.
arXiv preprint arXiv:2205.03318, 2022.

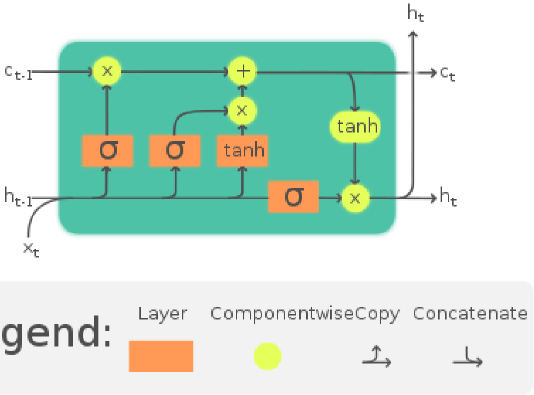
References III

Massimiliano Marcellino and Christian Schumacher.
Factor midas for nowcasting and forecasting with ragged-edge data: A model comparison for german gdp.
Oxford Bulletin of Economics and Statistics, 72(4):518–550, 2010.

James H Stock and Mark W Watson.
Forecasting output and inflation: The role of asset prices.
Journal of economic literature, 41(3):788–829, 2003.

James H Stock and Mark W Watson.
Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics.
In *Handbook of macroeconomics*, volume 2, pages 415–525. Elsevier, 2016.

LSTM Cell



Source: *(Wikipedia)*

Adaptive Momentum

$$m_w^{(t+1)} \leftarrow \beta_1 m_w^{(t)} + (1 - \beta_1) \nabla_w L^{(t)}$$

$$v_w^{(t+1)} \leftarrow \beta_2 v_w^{(t)} + (1 - \beta_2) \left(\nabla_w L^{(t)} \right)^2$$

$$\hat{m}_w = \frac{m_w^{(t+1)}}{1 - \beta_1^t}$$

$$\hat{v}_w = \frac{v_w^{(t+1)}}{1 - \beta_2^t}$$

$$w^{(t+1)} \leftarrow w^{(t)} - \eta \frac{\hat{m}_w}{\sqrt{\hat{v}_w} + \epsilon}$$

Source: *(Wikipedia)*