Finance Research Letters 61 (2024) 105003

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Fan charts in era of big data and learning☆

Jozef Baruník *,[1], Luboš Hanus[1,2]

Charles University, Czech Republic
Czech Academy of Sciences, Czech Republic

A R T I C L E I N F O

JEL classification:
C45
C53
E17
E37

Keywords:
Fan charts
Probabilistic forecasting
Machine learning
Deep learning
Macroeconomic time series

A B S T R A C T

We propose how to construct big data-driven macroeconomic fan charts, using machine learning methods to reflect the information in 216 relevant economic variables. Such data-rich fan charts do not rely on restrictive model assumptions and allow the exploration of non-Gaussian, asymmetric, heavy-tailed data and their non-linear interactions. By allowing complex patterns to be learned from a data-rich environment, our fan charts are useful for decision making that depends on the uncertainty of a potentially large number of economic variables — most public policy issues.

1. Introduction

A fundamental problem with economic variables is that they are inherently difficult to forecast (Stock and Watson, 2017). At the same time, most public policy issues involve uncertainty about future economic developments. Assessing and communicating forecast uncertainty has become a priority for central banks. As an early leader in recognizing the need to communicate the full forecast distribution rather than point estimates, the Bank of England started to communicate uncertainty with fan charts in 1996 (Britton et al., 1998). Over the following decades, most banks have used different methods to produce fan charts, with two leading approaches. First, some central banks rely on past forecast errors, assuming a normal distribution of risks and generating the central path using their core forecasting model. Second, other banks subjectively assess the uncertainty of future economic developments using a two-part normal distribution, assuming that the central forecast represents the mode of this distribution. An important stream of literature focuses on Bayesian fan charts (Cogley et al., 2005), which could be used to produce fan charts similar to those produced using expert judgement, but they would be model-based.

☆ We are grateful to the editor, Laura Ballester, and an anonymous referee for their useful comments and suggestions, which have greatly improved the paper. We are indebted to Wolfgang Hardle, Lukas Vacha, Martin Hronec, Frantisek Cech, and the participants at various conferences and research seminars for their valuable comments. We gratefully acknowledge the support from the Czech Science Foundation under the EXPRO GX19-28231X project. We provide the computational package `DistrNN.jl` in JULIA available at https://github.com/barunik/DistrNN.jl that allows one to obtain our measures on data the researcher desires.

* Correspondence to: Institute of Economic Studies, Charles University, Opletalova 26, 110 00, Prague, Czech Republic.
E-mail addresses: barunik@utia.cas.cz (J. Baruník), hanusl@utia.cas.cz (L. Hanus).
URL: https://barunik.github.io/ (J. Baruník).
[1] Institute of Information Theory and Automation, Czech Academy of Sciences, Pod Vodarenskou Vezi 4, 18200, Prague, Czech Republic.
[2] Institute of Economic Studies, Charles University, Opletalova 26, 110 00, Prague, CR and Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodarenskou Vezi 4, 18200, Prague, Czech Republic.

https://doi.org/10.1016/j.frl.2024.105003
Received 13 September 2023; Received in revised form 21 November 2023; Accepted 9 January 2024
Available online 11 January 2024

Worryingly, all these approaches rely on restrictive models and assumptions, do not take into account the (possibly non-linear) interaction of a large number of variables, nor well-known features of the data such as non-Gaussianity, asymmetries and heavy tails. In contrast, economists eager to use large numbers of series to understand fluctuations in economic data are collecting data unimaginable decades ago.[3] With the explosion in the volume, velocity and variety of data, the need to unlock the information hidden in big data has become a key topic in economics (Diebold, 2021). Challenged by the proliferation of parameters and strong criticism of arbitrarily chosen restrictions in both reduced and structured models in recent decades, economists wishing to explore the potentially rich information content of new datasets have recently turned their hopes to machine learning (Mullainathan and Spiess, 2017).

In this paper, we explore the use of machine learning for information-rich uncertainty forecasting of macroeconomic time series, and provide the data-driven fan charts in the era of big data. We construct a distributional machine learning method, based on recurrent neural network techniques, to provide probabilistic forecasts that reflect the time-series dynamics of potentially large amounts of available information. Our model selects the most preferable model from an unknown pool of models using innovative optimization techniques focusing on out-of-sample predictive performance and understanding the bias–variance trade-off.

It is important to note that we view probabilistic forecasting as a multivariate classification problem that maps the information set to a complete prediction distribution. Such a distributional regression approach is inversely related to approaches based on quantile regression (Wen et al., 2017). In contrast to our approach, quantile regression suffers from quantile crossing and thus provides potentially non-monotonous distribution predictions (Gasthaus et al., 2019).

2. (Deep) learning the fan charts

To be able to construct fan charts using machine learning, we introduce a multiple output neural network that will be used to approximate conditional distribution function. More formally, consider an economic time series y_t collected over $t = 1\dots,T$. Consider a partition of the support of y_t by $p > 1$ fixed thresholds corresponding to a set of empirical α_j quantiles $\{q^{\alpha_j}\}_{j=1}^{p}$, where $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_p < 1$ are p regularly spaced probability levels on a unit interval $[0,1]$. These partitions can be time-varying, so in general the elements of the partition are implicitly indexed by t.

The main goal then is to approximate a collection of conditional probabilities corresponding to the empirical quantiles such as

$$\Big\{F(q^{\alpha_1}),\dots,F(q^{\alpha_p})\Big\} = \Big\{\Pr\Big(y_{t+h} \le q^{\alpha_1}|\mathcal{I}_t\Big),\dots,\Pr\Big(y_{t+h} \le q^{\alpha_p}|\mathcal{I}_t\Big)\Big\} \tag{1}$$

for the collection of thresholds $1,\dots,p$, and use it for a h-step ahead probabilistic forecast made at time t with information $\mathcal{I}_t$ containing past values of y_t and possibly past values of other exogenous observable variables. Such probabilistic forecasts are usually highly dependent on model parametrization and quickly become infeasible as the number of covariates increases (Foresi and Peracchi, 1995; Anatolyev and Baruník, 2019). Stationarity of the data at hand is also a requirement that complicates forecasting as it is difficult to achieve in many cases. Our distribution neural network (DistrNN) aims to uncover non-linear and mostly complex relationships of time series without specifying a strict parametric structure and without requiring strict assumptions about the data, while focusing on the out-of-sample predictive power of the model.

We propose to approximate conditional distribution with an unknown general function. A set of probabilities that characterize conditional distribution function using set of predictors $z_t = (y_t, x_t^1,\dots,x_t^n)^\top$ is modelled jointly as

$$\Big\{\Pr\Big(y_{t+h} \le q^{\alpha_1}|z_t\Big),\dots,\Pr\Big(y_{t+h} \le q^{\alpha_p}|z_t\Big)\Big\} = \mathfrak{g}_{W,b}(z_t), \tag{2}$$

where $\mathfrak{g}_{W,b}$ is a multiple output neural network with L hidden layers that we name as distributional neural network:

$$\mathfrak{g}_{W,b}(z_t) = g^{(L)}_{W^{(L)},b^{(L)}} \circ \dots \circ g^{(1)}_{W^{(1)},b^{(1)}}\left(z_t\right), \tag{3}$$

where $W = \left(W^{(1)},\dots,W^{(L)}\right)$ and $b = \left(b^{(1)},\dots,b^{(L)}\right)$ are weight matrices and bias vector. Any weight matrix $W^{(\ell)} \in \mathbb{R}^{m\times n}$ contain m neurons as n column vectors $W^{(\ell)} = [w^{(\ell)}_{\cdot,1},\dots,w^{(\ell)}_{\cdot,n}]$, and $b^{(\ell)}$ are thresholds or activation levels which contribute to the output of a hidden layer allowing the function to be shifted.

Variables that we consider evolve over time, and hence traditional neural networks assuming independence of data may not approximate relationships sufficiently well. Recurrent neural networks (RNN) are more suitable for many economic problems. Particularly, we consider LSTM that incorporates memory units into the structure (Hochreiter and Schmidhuber, 1997) and capture potentially long time dynamics in the time series. Figure 1 illustrates the model.

2.1. Loss function

To represent the fan charts, we need the cumulative distribution function that is a non-decreasing function bounded on $[0,1]$. The problem is essentially a more complex classification closely related to logistic regression, hence we use a binary cross-entropy loss function, and we introduce a penalty to order the predicted probabilities:

[3] A century ago, in the 1920s, the Harvard Economic Service provided economic indices and forecasts based on available data (Friedman, 2009). Later, 1277 time series were used to study business cycles by Lerner (1947).

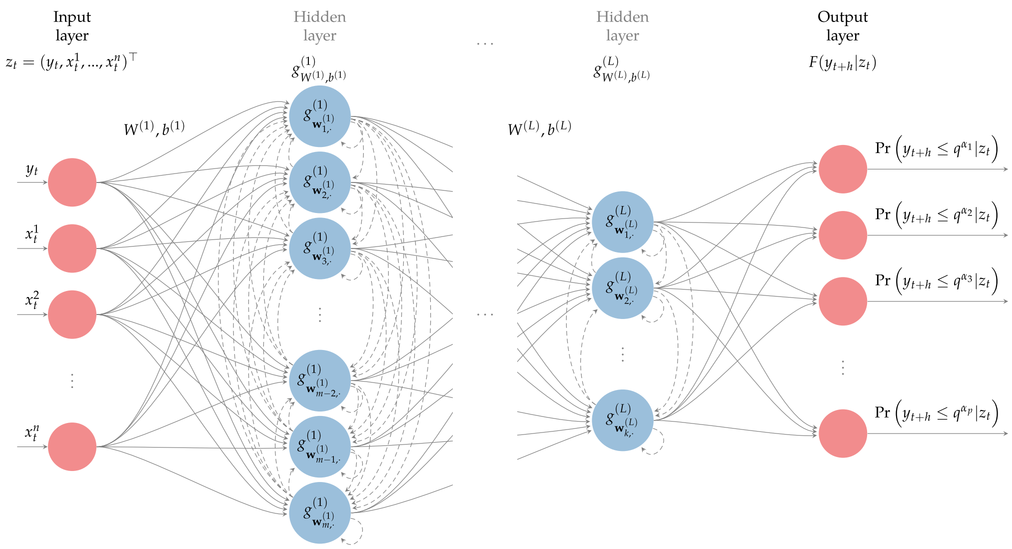

Fig. 1. Distributional (Deep) recurrent network.
An illustration of a deep recurrent neural network $\mathfrak{g}_{W,b}(z_t)$ that captures relationship between all nodes (solid) and recurrent paths (dashed) in the network at time t to model the collection of conditional probabilities $\left\{\Pr\left(y_{t+h} \leq q^{\alpha_1}|z_t\right), \ldots, \Pr\left(y_{t+h} \leq q^{\alpha_p}|z_t\right)\right\}$ with set of predictor variables $z_t = (y_t, x_t^1, \ldots, x_t^n)^\top$. With large number of hidden layers L the network is deep.

$$\mathcal{L} = \underbrace{-\frac{1}{T}\sum_{t}^{T}\frac{1}{p}\sum_{j}^{p}\left(\mathbb{I}_{\{y_{t+h}\leq q^{\alpha_j}\}}\log\left\{\widehat{\mathfrak{g}}_{W,b,j}(z_t)\right\} + \left(1-\mathbb{I}_{\{y_{t+h}\leq q^{\alpha_j}\}}\right)\log\left\{1-\widehat{\mathfrak{g}}_{W,b,j}(z_t)\right\}\right)}_{\text{binary cross-entropy}}$$

$$+ \underbrace{\lambda_m \sum_{t}^{T}\sum_{j=1}^{p-1}\left(\widehat{\mathfrak{g}}_{W,b,j}(z_t) - \widehat{\mathfrak{g}}_{W,b,j+1}(z_t)\right)_+}_{\text{monotonicity penalty}} \tag{4}$$

where $(u)_+$ is a rectified linear units function, ReLU, $(u)_+ = \max\{u, 0\}$, which passes through only positive differences between two neighbouring values, j and $j+1$, of CDF, those violating the monotonicity condition, and $\mathbb{I}_{\{.\}}$ is an indicator function. This violation is controlled by the penalty parameter λ_m. Note that in addition to its simplicity, ReLU is used for convenience reasons allowing for general use.

2.2. Learning, regularization and hyper-parameters

Due to the high dimensionality and non-linearity of the problem, estimation of a deep neural network is a complex task. Selection of hyper-parameters together with regularization methods play crucial role in reduction of risk from estimation. In particular, we use ReLU activation function to introduce non-linearity to our problem, and help the optimization algorithm converge faster. For the learning process, we use adaptive gradient algorithm, AdamW (Kingma and Ba, 2014; Loshchilov and Hutter, 2019).

We use hyper-optimization algorithms based on random search over a grid/cube of parameter ranges. Using hyper-optimization[4] we select the learning rate of the optimizer, η, the weight decay parameter of AdamW, λ_W, and a Dropout parameter regularizing models (Srivastava et al., 2014) that is an efficient way of performing model averaging with neural networks. We also use the early stopping technique that helps regularization to prevent from over-fitting.

3. Macroeconomic fan charts in era of big data

We construct a data-driven macroeconomic fan chart from a best approximation model learned from hundreds of variables using deep learning. This is in sharp contrast to the literature that provides the uncertainty of macroeconomic variables using so-called predictive fan charts (Britton et al., 1998; Stock and Watson, 2017) using a few variables with a parameterized and structured model that requires a number of assumptions. To benchmark our approach, we consider a state-of-the-art Bayesian fan charts using Bayesian Vector Autoregression (BVAR) model with factor components (McCracken and Ng, 2020) that are widely used in the macroeconomic literature. In addition, we compare it to the state-of-the art in the neural network time series probabilistic forecasting literature, DeepAR (Salinas et al., 2020).

[4] We using Julia package HyperOpt.jl (https://github.com/baggepinnen/Hyperopt.jl).

Table 1
Fan chart recurrent DistrNN parameters space for the empirical application.

Fixed parameters	Value	Hyper parameters	Values
Number of layers	2	Learning rate, η	0.0001, 0.001, 0.005
Mini batch size	8	Dropout rate, ϕ	0.2, 0.4
Epochs	350	L_2-decay regularization rate, λ_W	0.00001, 0.00005
Monotonicity parameter, λ_m	5.0	Nodes dimensions	32 × 32, 64 × 64, 60 × 50
Cross-validation, k-folds	3		
Train/validation ratio	0.93		

The hyperoptimization algorithm searches through the whole hyperparameter space and tries all sets/combinations of hyperparameters to evaluate the model.

3.1. Data

To construct such a tool for measuring uncertainty, we use a high-dimensional dataset of McCracken and Ng (2020), which has been widely used in the macroeconomic literature (Coulombe et al., 2022) and is available on the website of the Federal Reserve of St-Louis. Our dataset contains 216 quarterly US macroeconomic and financial indicators, observed from 1961Q1 to 2019Q4. Since the number of variables is non-stationary, we follow the transformation codes used by McCracken and Ng (2020). Using this dataset, we construct a data-rich fan chart for real GDP growth (GDPC1), inflation (CPIAUCSL) and unemployment rate (UNRATE), which, to the best of our knowledge, will be the first of its kind to reflect high-dimensional information from 216 relevant variables.

3.2. Deep-learning based fan charts

To obtain a h-step ahead forecast that forms a fan chart, we consider a direct forecasting scheme. Exploring the data structures, we form h distribution networks $\hat{\mathfrak{g}}^{(1)}_{W,b}(z_t), \dots, \hat{\mathfrak{g}}^{(h)}_{W,b}(z_t)$, where the entire (continuous) h step-ahead conditional distribution $\hat{\mathfrak{g}}^{(h)}_{W,b}(z_t)$ is obtained by interpolating the cumulative distribution function, preserving the monotonicity of the result. Here we apply the Fritsch–Carlson monotonic cubic interpolation (Fritsch and Carlson, 1980), and use the predicted cumulative distribution function $\hat{F}_{t+h}(\alpha|\mathcal{I}_t)$ to form $k - size$ prediction intervals for a fan chart as

$$PI^{k}_{t+h} = \left[\hat{F}^{-1}_{t+h}(\alpha_l|\mathcal{I}_t), \hat{F}^{-1}_{t+h}(\alpha_u|\mathcal{I}_t)\right], \tag{5}$$

such that $k = \alpha_u - \alpha_l$ is size of the interval.

To benchmark our approach to the existing approaches, we compare the predictions with the state-of-the-art Bayesian vector autoregression estimated on the factors extracted from the data as in McCracken and Ng (2020), and neural network base DeepAR (Salinas et al., 2020) with external variables.[5] We evaluate the h-step-ahead forecasts with quantile loss function (Clements et al., 2008)

$$L^{h}_{\alpha,m} = E\left[(\alpha - \mathbb{I}\{e_{t+h,m} < 0\})e_{t+h,m}\right], \tag{6}$$

for a model m, horizon h, and α-quantile where $e_{t+h,m} = y_{t+h} - \hat{F}^{-1}_{t+h,m}(\alpha|\mathcal{I}_t)$ is the difference between the original time series and α-quantile forecast given the information set, $\mathcal{I}_t$.

3.3. Setup

Working with quarterly data, we compute $h = 1, \dots, 6$ horizon forecasts for each quarter of the out-of-sample period, starting with 2012:Q3 and ending with 2019:Q4. The conditional distribution is approximated using $j = 1, \dots, 19$ empirical $\alpha_j = (0.01, \dots, 0.99)$ probability levels. The learning explores 36 combinations of hyperparameters to find the best approximating model for each h-step ahead forecast separately. The hyperparameter space is optimized once on the training and validation parts prior out-of-sample, and the training procedure performs a growing window forward validation scheme on the training data using 3-fold cross-validation. We standardize the data with respect to training and validation in order not to contaminate our training data with information from the out-of-sample part.

Table 1 summarizes all parameters and details used in the estimation. The choice of parameters such as mini batch size, size of network, k-folds are based on time dimension sample size. While the choices may seem arbitrary, we have experimented with large number of other setups, and finally used the models that deliver balanced trade-off between computational time and precision.

3.4. Discussion

We start the discussion by presenting the qualitative results of the GDP growth, inflation and unemployment forecasts in the form of fan charts. Fig. 2 compares the median as well as the 50% and 90% prediction intervals over four different periods.

[5] We use GluonTS (Alexandrov et al., 2020) implementation and included all macroeconomic variables as external time dynamic features. For more information see https://ts.gluon.ai/stable/tutorials/forecasting/extended_tutorial.html.

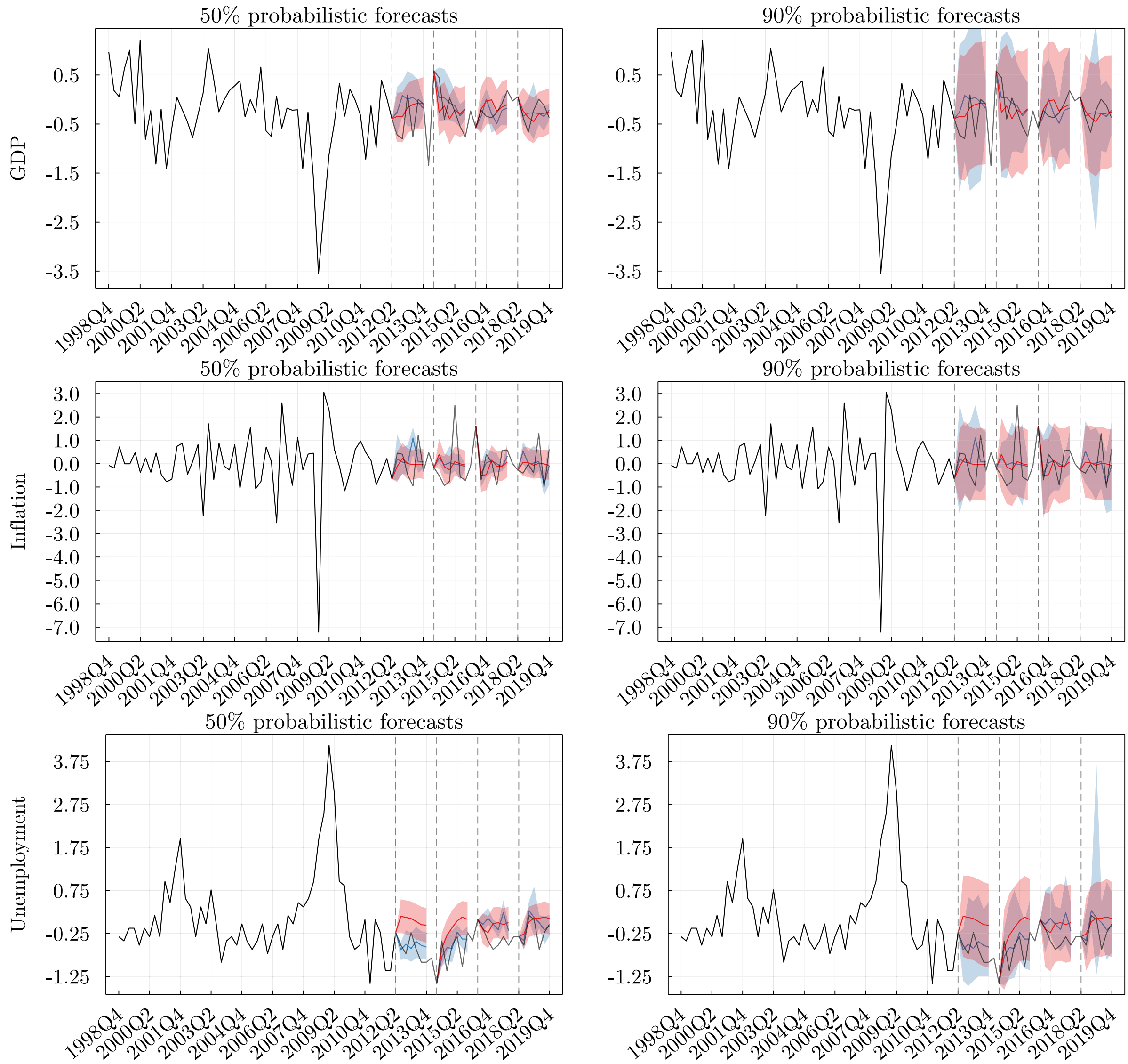

Fig. 2. Deep-learning based (blue) and BVAR (red) fan charts.
6-step-ahead quarterly forecasts of GDP growth (top), Inflation (middle), and Unemployment rate (bottom) with 50% (left column) and 90% (right column) fan charts obtained by distributional network (blue) using 216 quarterly US macroeconomic and financial indicators from the FRED-QD database, and a factor three+four-variable BVAR (red). Forecasts are made at the end of the 2012:Q2, 2014:Q2, 2016:Q2 and 2018:Q2 depicted by dashed vertical lines. Train data are plotted by black solid line and test data by grey solid line.

Prediction intervals from distributional neural networks are asymmetric and, unlike traditional time series represented by BVAR, are not very smooth over the forecast horizon. With increasing uncertainty about the future, DistrNN still learns some structure from the data and the probability intervals are less similar to the shape of "fans". Asymmetry in the forecasts is an important feature for a policy maker because it reflects the uneven occurrence of events in probability. Put simply, such fan charts signal to the user an uneven (or unbalanced) distribution of data around a central point. The intervals are narrower compared to the BVAR in most cases, especially when looking at the 50% intervals. In the case of the unemployment rate, the BVAR model is less able to reduce the uncertainty about the future observation, probably because of strong peaks in previous years. In contrast, our DistrNN captures the uncertainty well.

While Fig. 2 is illustrative, it only shows few periods and to support the gains of deep learning approach we further quantify the prediction differences for whole out-of-sample period. Table 2 presents the quantitative comparison of predictions. We compare the forecasts at $h = 1, \ldots, 6$ horizons using tick loss (Eq. (6)) for selected $\alpha = \{0.1, 0.25, 0.5, 0.75, 0.9\}$ probability levels.

In addition to the quantile losses of the DistrNN and BVAR forecasts, Table 2 shows the DeepAR losses. Our deep learning based DistrNN approach provides forecasts with lower error (in blue) at most of the probability levels and horizons considered, and

Table 2
Quantile loss of DistrNN and BVAR.

	0.1			0.25			0.5			0.75			0.9		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
GDP															
$h = 1$	2.395	*2.389*	4.480	**3.351**	3.531	4.888	**4.472**	4.717	4.802	**3.659**	4.666	4.072	**2.170**	2.903	2.934
$h = 2$	**2.224**	2.248	4.514	3.832	*3.811*	4.936	4.553	*4.362*	5.132	**3.397**	4.361	4.757	**2.184**	2.676	4.129
$h = 3$	3.089	*2.604*	3.948	4.253	*4.082*	4.364	4.564	5.013	*4.509*	**3.980**	4.379	4.102	**2.560**	2.825	3.449
$h = 4$	**2.053**	2.169	4.892	3.661	*3.609*	5.340	**4.336**	4.526	5.412	**3.298**	4.311	4.951	**1.950**	2.833	4.169
$h = 5$	**1.962**	2.324	4.050	**3.246**	3.708	4.656	**3.783**	4.203	5.223	**2.573**	4.265	5.141	**1.595**	2.703	4.707
$h = 6$	2.176	*2.159*	4.072	3.298	*3.036*	4.619	**3.776**	4.070	4.929	**2.682**	4.307	4.723	**1.909**	2.733	4.195
Inflation															
$h = 1$	4.902	*3.018*	4.481	6.981	5.823	*5.445*	7.574	7.968	*6.292*	7.151	*6.374*	6.622	4.443	*3.547*	6.462
$h = 2$	3.534	*3.070*	5.598	6.387	*4.998*	6.279	8.176	8.031	*6.667*	6.857	7.436	*6.414*	**3.815**	4.331	5.844
$h = 3$	**2.405**	3.059	4.417	5.166	*4.948*	5.444	8.329	8.034	*6.215*	8.579	7.703	*6.366*	5.926	*4.939*	6.101
$h = 4$	3.420	*3.021*	4.348	6.863	*5.234*	5.308	9.330	8.251	*6.218*	8.195	7.901	*6.568*	5.207	*4.778*	6.404
$h = 5$	3.572	*3.097*	6.571	6.149	*5.332*	7.552	8.199	8.274	*8.154*	**7.035**	7.817	8.044	**4.525**	4.809	7.502
$h = 6$	4.004	*3.055*	5.453	6.747	*5.215*	6.513	8.198	7.840	*7.153*	**6.941**	7.376	7.060	**4.645**	4.866	6.523
Unemployment															
$h = 1$	**1.664**	1.686	4.416	**2.871**	3.489	4.556	**3.489**	4.959	4.092	**2.891**	4.203	3.184	**1.739**	2.516	2.412
$h = 2$	2.001	*1.391*	4.681	**3.150**	3.402	5.051	**3.831**	5.152	4.895	**3.081**	4.456	4.294	**1.805**	2.635	3.614
$h = 3$	1.925	*1.728*	5.772	3.432	*3.231*	5.925	**4.039**	4.512	5.313	**3.528**	4.132	4.135	**2.408**	2.493	3.134
$h = 4$	2.564	*1.674*	6.328	**3.569**	3.857	6.210	**4.036**	5.506	5.193	**3.508**	4.481	3.714	**2.313**	2.626	2.543
$h = 5$	**2.132**	2.452	8.201	**3.793**	4.244	8.053	**4.561**	5.797	6.727	**3.305**	5.007	4.851	**1.780**	2.676	3.380
$h = 6$	2.113	*1.916*	8.644	**3.253**	3.751	8.266	**3.731**	5.741	6.796	**2.929**	4.576	4.835	**1.756**	2.599	3.424

Note: Quantiles losses of Distributional Recurrent Neural Network (1), Bayesian VAR (2), and DeepAR (3), for variables GDP growth (GDP), Inflation, and Unemployment rate. The out-of-sample forecasts for 25 quarters are made at α-levels $\{0.1, 0.25, 0.5, 0.75, 0.9\}$, horizons $h = 1, \dots, 6$, starting at Q3/2012 and ending at Q4/2019. Cases with DistrNN forecast loss being smaller are blue and bold, when a benchmark loss is lowest the colour is green and italic.

provides greater improvement at shorter horizons. Notable gains are at 75% and 90%, where DistrNN dominates strongly. With the exception of inflation, DistrNN also improves losses at the median and 25% levels.

4. Conclusion

In this paper, we have proposed a new approach to constructing fan charts using large number of economic variables and state-of-the-art machine learning methods. The distributional neural network relaxes the assumption on the distributional family of time series and allows the model to fully explore the data. The approach is particularly useful for modelling data with non-Gaussian, non-linear and asymmetric structures. We have shown that our distributional neural network is useful in constructing big data-driven macroeconomic fan charts, which are the first of their kind as they are learned from the structure between 216 relevant economic variables.

CRediT authorship contribution statement

Jozef Baruník: Conceptualization, Methodology, Writing – original draft, Writing – review & editing. **Luboš Hanus:** Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

Alexandrov, A., Benidis, K., Bohlke-Schneider, M., Flunkert, V., Gasthaus, J., Januschowski, T., Maddix, D.C., Rangapuram, S., Salinas, D., Schulz, J., Stella, L., Türkmen, A.C., Wang, Y., 2020. GluonTS: Probabilistic and Neural Time Series Modeling in Python. J. Mach. Learn. Res. 21 (116), 1–6.
Anatolyev, S., Baruník, J., 2019. Forecasting dynamic return distributions based on ordered binary choice. Int. J. Forecast. 35 (3), 823–835.
Britton, E., Fisher, P., Whitley, J., 1998. The inflation report projections: understanding the fan chart. Chart 8 (10).
Clements, M.P., Galvão, A.B., Kim, J.H., 2008. Quantile forecasts of daily exchange rate returns from forecasts of realized volatility. J. Empir. Financ. 15 (4), 729–750.

Cogley, T., Morozov, S., Sargent, T.J., 2005. Bayesian fan charts for UK inflation: Forecasting and sources of uncertainty in an evolving monetary system. J. Econom. Dynam. Control 29 (11), 1893–1925.
Coulombe, P.G., Leroux, M., Stevanovic, D., Surprenant, S., 2022. How is machine learning useful for macroeconomic forecasting?. J. Appl. Econometrics 37 (5), 920–964.
Diebold, F.X., 2021. What's the big idea? Big data and its origins. Significance 19, 36–37.
Foresi, S., Peracchi, F., 1995. The conditional distribution of excess returns: An empirical analysis. J. Amer. Statist. Assoc. 90 (430), 451–466.
Friedman, W.A., 2009. The harvard economic service and the problems of forecasting. History Polit. Econ. 41 (1), 57–88.
Fritsch, F., Carlson, R., 1980. Monotone piecewise cubic interpolation. SIAM J. Numer. Anal. 17 (2).
Gasthaus, J., Benidis, K., Wang, Y., Rangapuram, S.S., Salinas, D., Flunkert, V., Januschowski, T., 2019. Probabilistic forecasting with spline quantile function rnns. In: The 22nd International Conference on Artificial Intelligence and Statistics. pp. 1901–1910.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.
Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
Lerner, A.P., 1947. Measuring business cycles. By Arthur F. Burns and Wesley C. Mitchell.[Studies in Business Cycles, Vol. II.] New York: National Bureau of Economic Research, 1946. pp. xxvii, 560. J. Econ. History 7 (2), 222–226.
Loshchilov, I., Hutter, F., 2019. Decoupled weight decay regularization. arXiv preprint arXiv:1711.05101.
McCracken, M., Ng, S., 2020. FRED-QD: A quarterly database for macroeconomic research. Technical report, National Bureau of Economic Research NBER Working Paper No. 26872.
Mullainathan, S., Spiess, J., 2017. Machine learning: an applied econometric approach. J. Econ. Perspect. 31 (2), 87–106.
Salinas, D., Flunkert, V., Gasthaus, J., Januschowski, T., 2020. DeepAR: Probabilistic forecasting with autoregressive recurrent networks. Int. J. Forecast. 36 (3), 1181–1191.
Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15 (1), 1929–1958.
Stock, J.H., Watson, M.W., 2017. Twenty years of time series econometrics in ten pictures. J. Econ. Perspect. 31 (2), 59–86.
Wen, R., Torkkola, K., Narayanaswamy, B., Madeka, D., 2017. A multi-horizon quantile recurrent forecaster. arXiv preprint arXiv:1711.11053.