Macrocasting

Tanujit Chakraborty

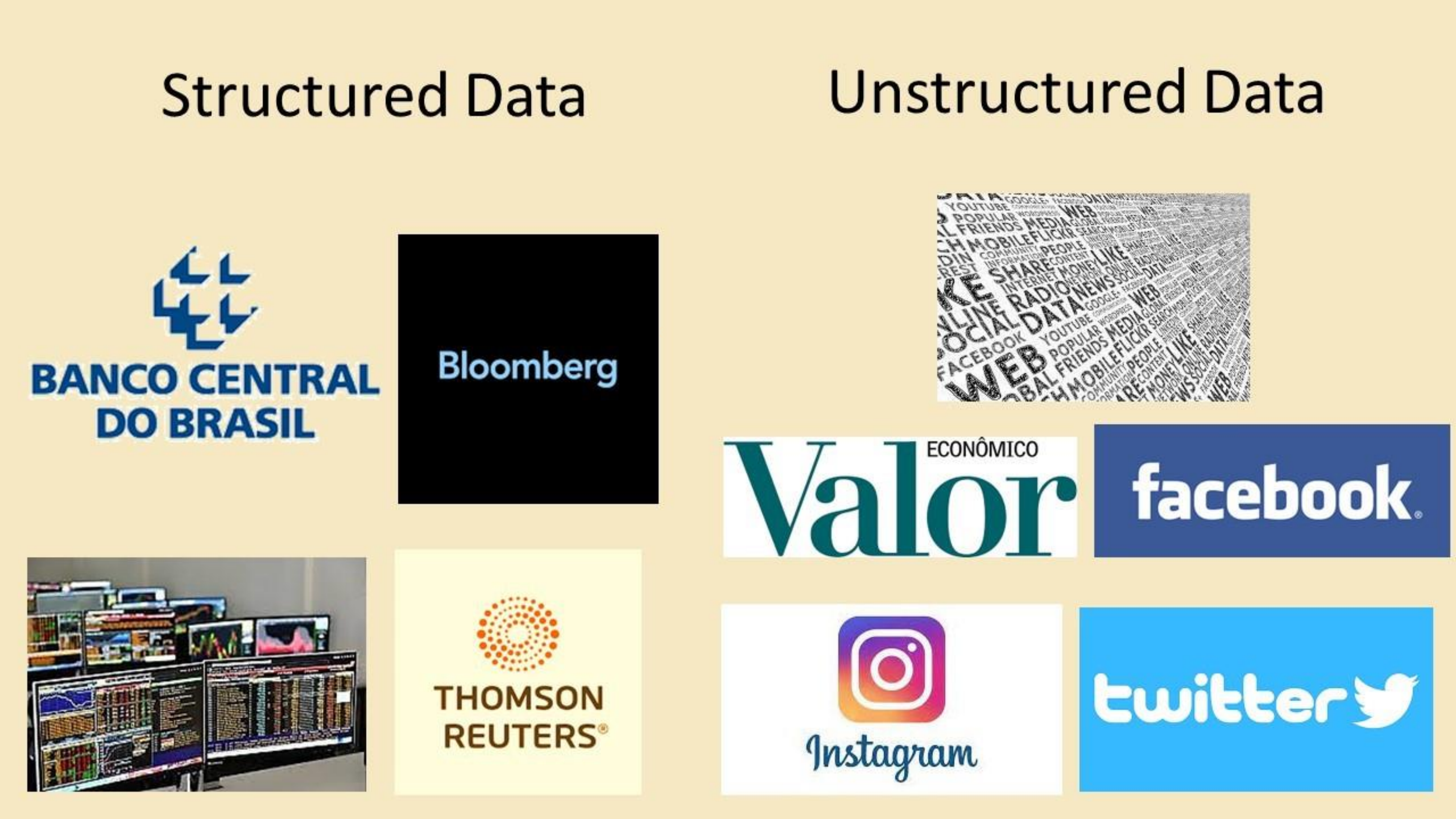
Associate Professor of Statistics & Data Science

Sorbonne University Abu Dhabi

Big Data in Economics and Finance

- "Big Data" boom in Economics and Finance
  - Structured and unstructured data.
- Standard econometric methods fail to handle such data.
  - Curse of dimensionality; post model-selection inference (worse with big data).
- Solution: New high-dimensional statistical and machine learning methods.
  - –Penalized regressions, (nonlinear) factor models, ensemble, tree-based models, neural-networks, ...
  - –Natural language processing methods.

Big Data in Economics and Finance



Big Data in Economics and Finance

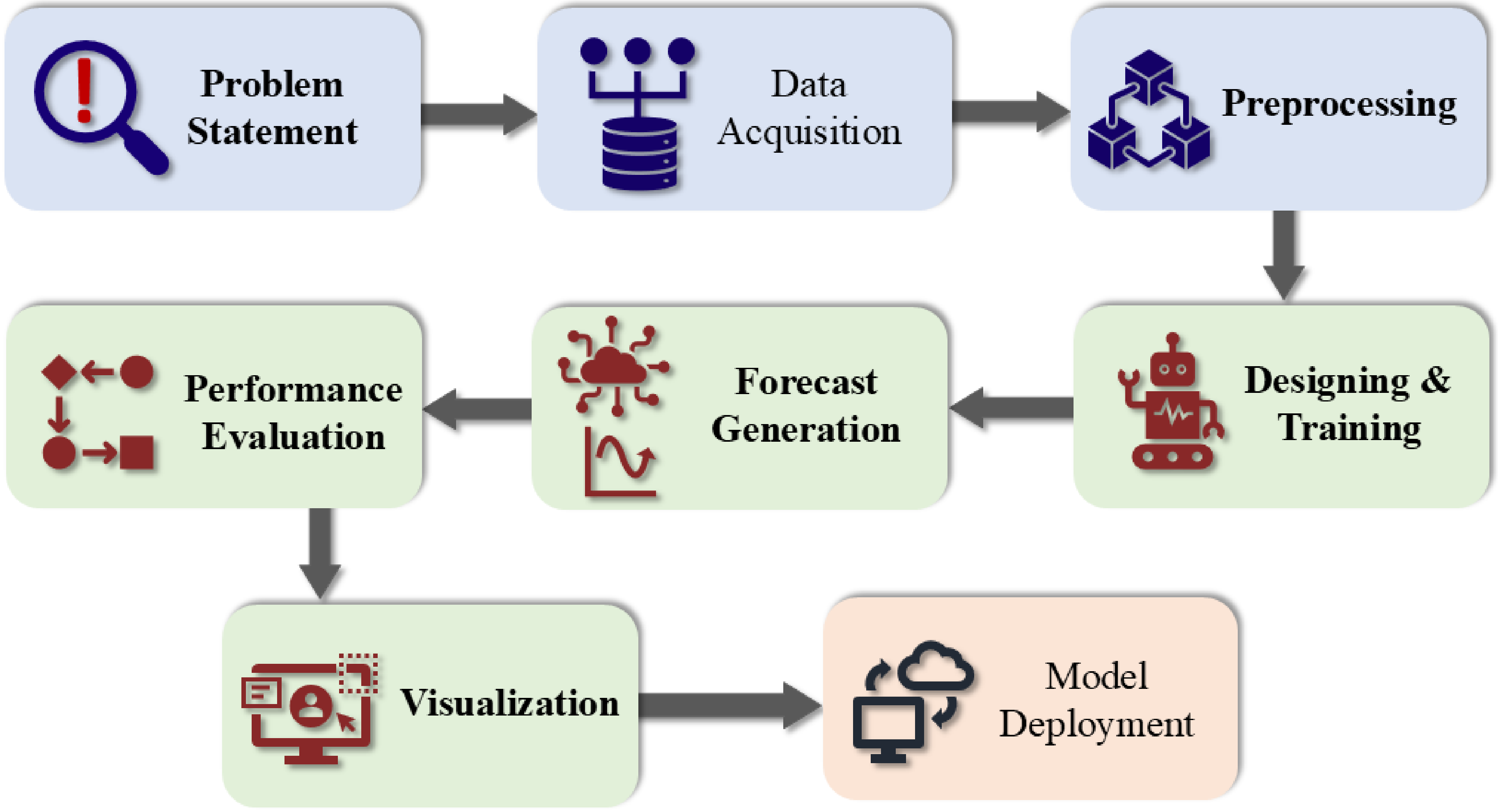


Figure: Business forecasting workflow with main strata: data (blue rectangles), analytics (green rectangles), and deployment (orange rectangles).

Tools for Data Scientists

- Statistics is the study of the collection, analysis, interpretation, presentation and organization of data.
- Data science is the study of the generalizable extraction of knowledge from data, yet the key word is science.
- Machine learning is the sub-field of computer science that gives computers the ability to learn without being explicitly programmed.
- Artificial Intelligence research is defined as the study of intelligent agents: any device that perceives its environment and takes actions that maximize its chance of success at some goal.
- Forecasting is estimating how the sequence of observations will continue into the future. (e.g., Forecasting of major economic variables like GDP, Unemployment, Inflation, Exchange rates, Production and Consumption).

What people think I forecast?

When I go to any university, and I tell people that I work on time series forecasting and machine learning, usually one of two things happens:

- ...like, weather forecasting?
  - Lots of domain knowledge and specialized models exist
  - We leave it to meteorologists
- ...so, can you predict the stock market and we all get rich?
  - I'll tell you how, and we're all going to be rich!
  - Try it on your own risk!

What I forecast?

- Key macroeconomic variables (inflation, unemployment, exchange rate, etc.)
- Epidemic time series (e.g., dengue, malaria, hepatitis, etc.)
- Sales forecasting in the supply chain, retail at pharmacy companies
- Forecasting in climate
  - Air quality
  - El Nino
  - Seismic events
- ...

Can these be forecasted? (Easy/Hard)

1. Daily electricity demand in 3 days' time
2. Google stock price tomorrow
3. Google stock price in 6 months' time
4. Maximum temperature tomorrow
5. Total sales of drugs in pharmacies next month

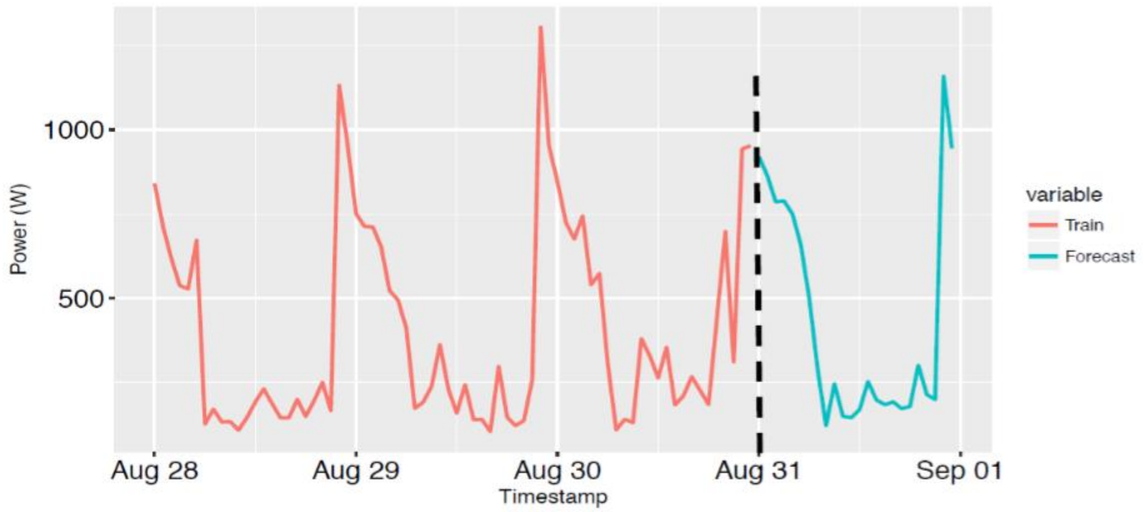
Something is easy to forecast if:

1. We have a good understanding of the factors that contribute to it

2. There is a lot of data available

3. The future is somewhat similar to the past

    - ID assumption: samples are identically distributed



4. The forecasts cannot affect the thing we are trying to forecast.

    - self-fulfilling prophecies (election polls)
    - controlled systems
    - Big bull effect in stock markets / bitcoin prices

Forecasters are to blame!

- News report on 16 August 2006: A Russian woman is suing weather forecasters for wrecking her holiday. A court in Uljanovsk heard that Alyona Gabitova had been promised 28 degrees and sunshine when she planned a camping trip to a local nature reserve, newspaper Nowyje Iswestija said.

- But it did nothing but pour with rain the whole time, leaving her with a cold. Gabitova has asked the court to order the weather service to pay the cost of her travel.

Reputations can be made and lost

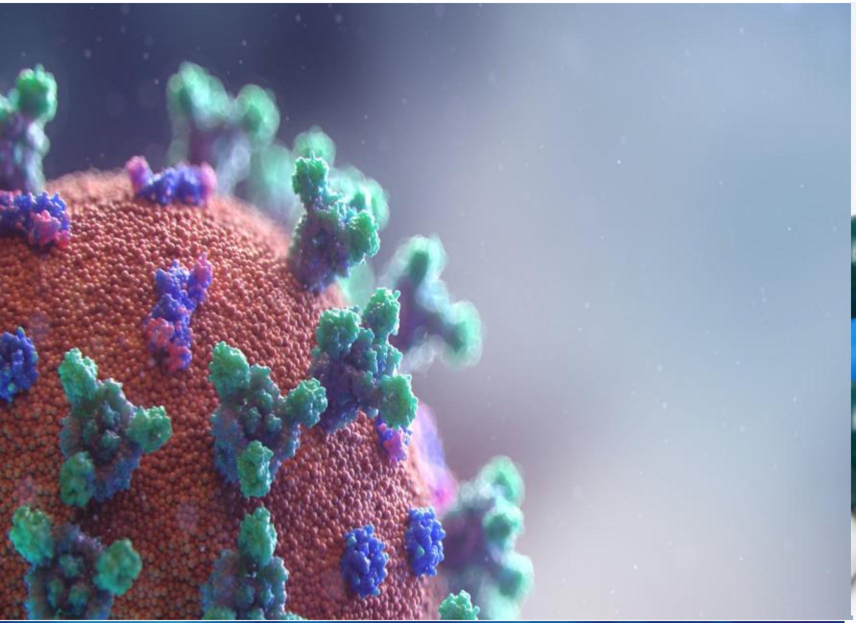
Some Misconceptions (Low Expectations): Our forecasts will always be inaccurate, so we should concentrate our efforts elsewhere.

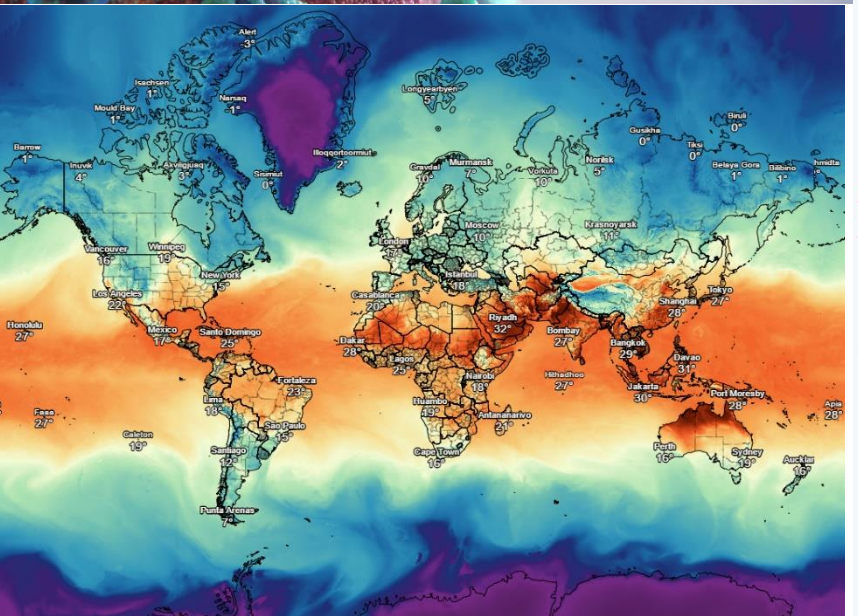
"There's no chance that the iPhone is going to get any significant market share. No chance."
(Steve Ballmer, CEO Microsoft, April 2007)

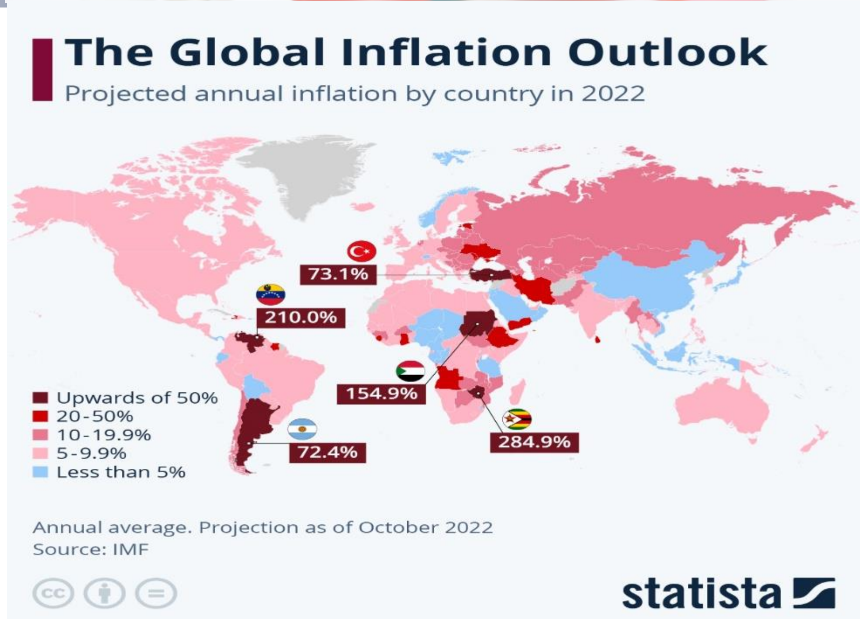
Some Misconceptions (High Expectations): If only we had the latest forecasting technology, then all our problems could be solved.

"We're going to be opening relatively soon … The virus … will go away in April."
(Donald Trump, February 2020)

What can we forecast?



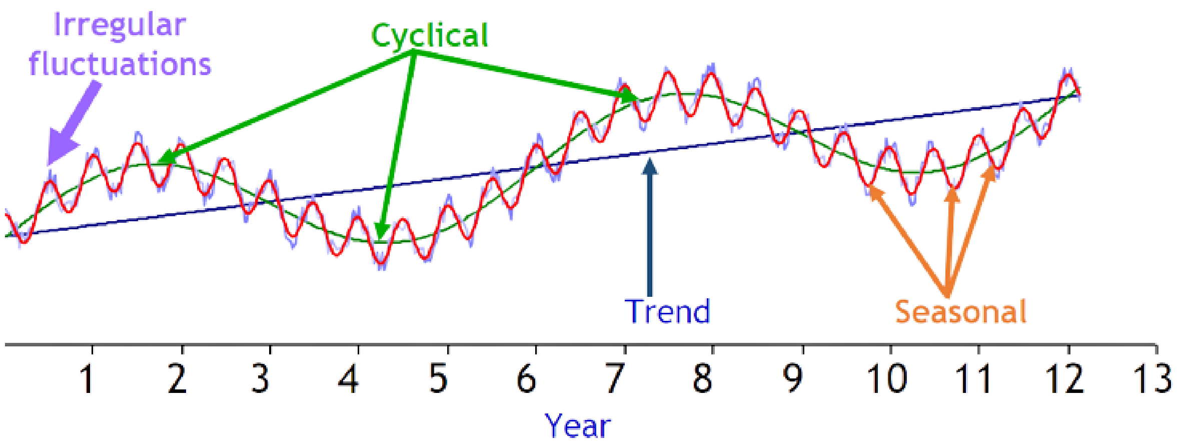


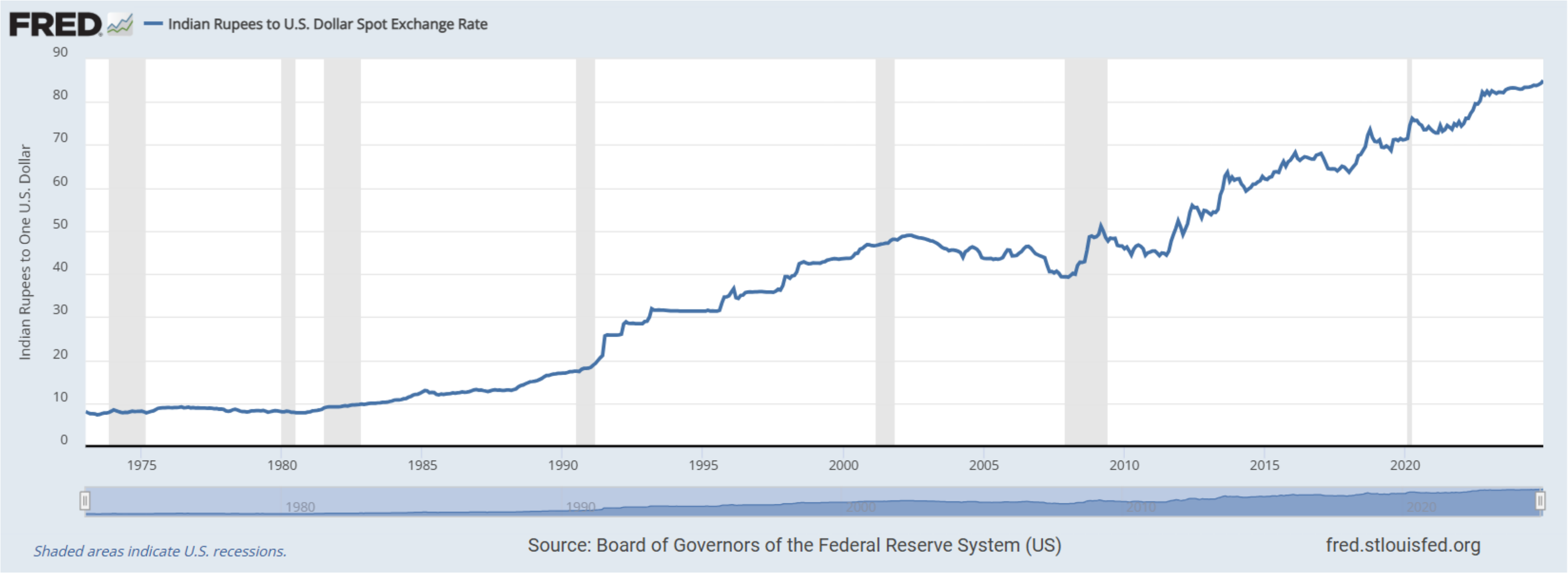


What is time series?

Time series is a set of observations, each one being recorded at a specific time.

Stationary time series is roughly horizontal, constant variance and no patterns predictable in the long-term.

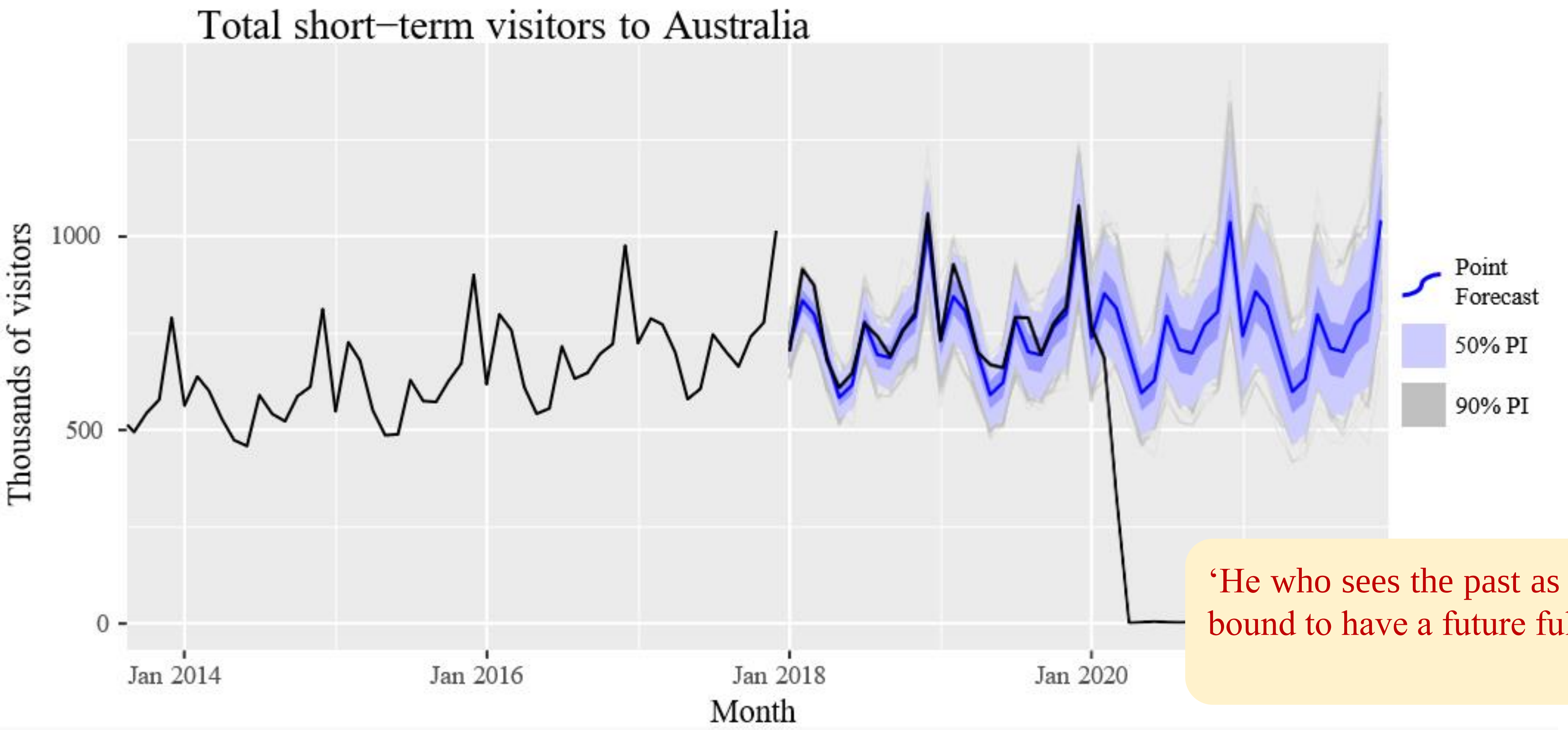




Units : Indian Rupees to One U.S. Dollar, Not Seasonally Adjusted
Frequency : Monthly (Averages of daily figures)

What is a forecast?

Forecasting is estimating how the sequence of observations will continue into the future.



'He who sees the past as surprise-free is bound to have a future full of surprise.'
- Amos Tversky

Forecasting approaches

Types of forecasting approaches

Judgemental

Role playing

Expert opinion (Delphi method)

Scientific

Statistical Methods (ES, ARIMA, GP)

ML Methods (LSTM, ARNN)

JOURNAL OF ECONOMIC SURVEYS

ARTICLE | Full Access

EVALUATING MACROECONOMIC FORECASTS: A CONCISE REVIEW OF SOME RECENT DEVELOPMENTS

Philip Hans Franses, Michael McAleer, Rianne Legerstee

ANNUAL REVIEW OF ECONOMICS Volume 10, 2018

Macroeconomic Nowcasting and Forecasting with Big Data

Brandyn Bok[1], Daniele Caratelli[2], Domenico Giannone[1], Argia M. Sbordone[1], and Andrea Tambalotti[1]

Advantages and Disadvantages of Statistical Models

1. Based on empirical data
2. Objective measure of uncertainty.
3. Computable, Replicable, Testable.
4. Able to compute prediction intervals.

- Extrapolate trend, seasonal, and auto-correlation effect into the near future (relies on explanatory variables)
- Assumptions don't always hold (seasonality, stationarity!)
- Needs historical data, less suitable for novel (e.g., pandemic) situations

Advantages and Disadvantages of ML-DL Models

1. Smart algorithms, few assumptions and applied to huge data sets.
2. Solve problems which traditional statistical methods can’t handle (largely due to size of data sets).
3. Strong emphasis on out-of-sample predictive performance (the test data).

- Often computationally expensive, low-test accuracy
- Less interpretable (e.g., estimating the effect of explanatory variable)

- More recently: Data assimilation methods (e.g., Kalman Filters, Particle Filtering) and Time frequency domain tools from signal image processing (e.g., Wavelet and Fourier decomposition)

- Finally – assess forecast accuracy on test set and assess the impact the accuracy is having.

Macrocasting using FEWNet

ELSEVIER

International Journal of Forecasting

Available online 20 September 2024

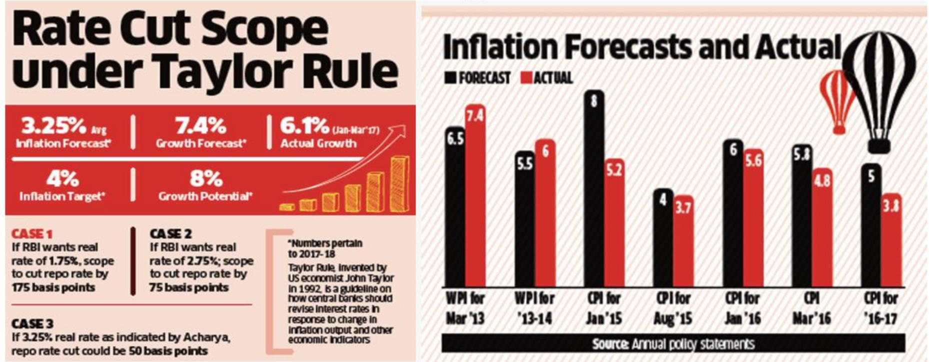
In Press, Corrected Proof What's this?

Forecasting CPI inflation under economic policy and geopolitical uncertainties ☆

Shovon Sengupta [a,b,1], Tanujit Chakraborty [c,d,1], Sunny Kumar Singh [b]

Motivation

Why Reserve Bank of India should fix inflation forecasting model now

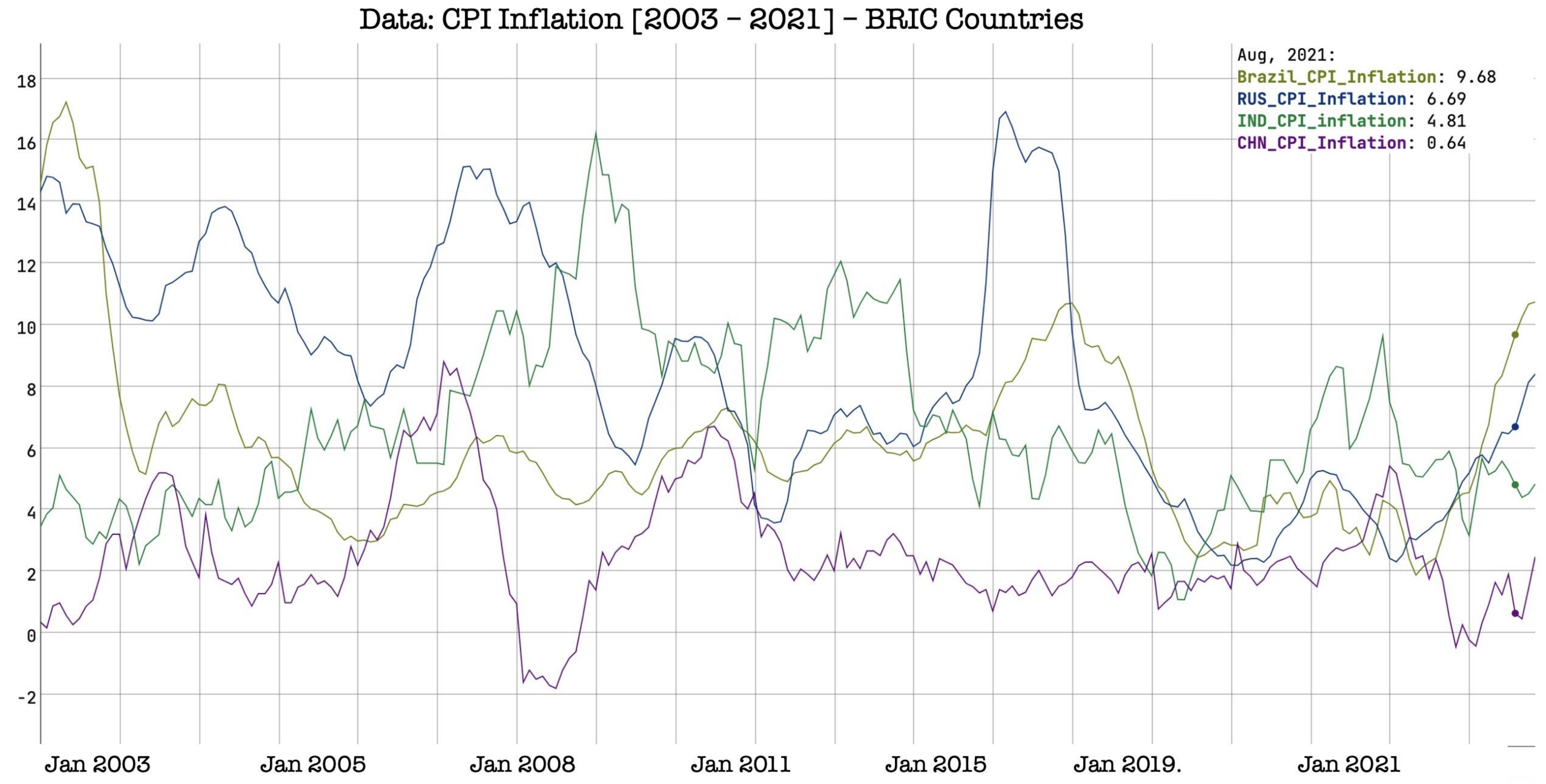


ECB's huge forecasting errors undermine credibility of current forecasts

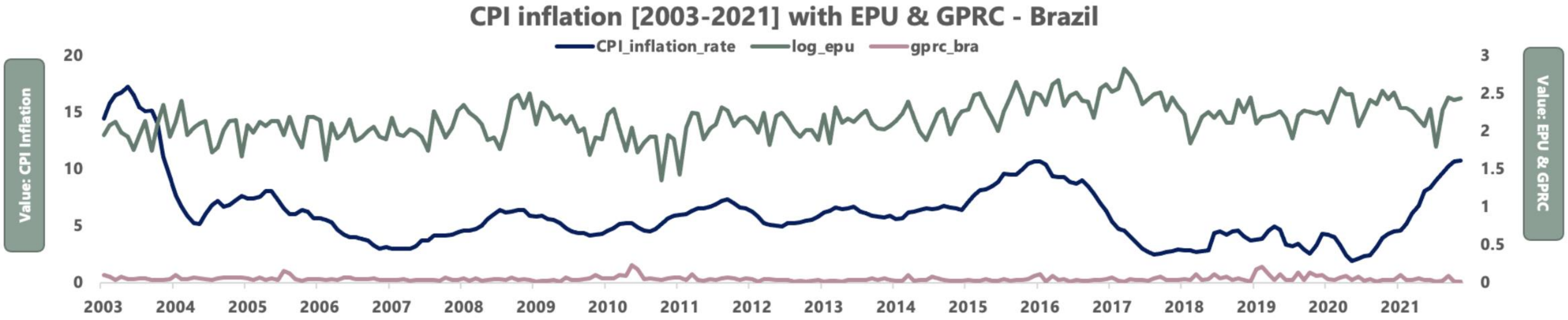
In the past five years ECB forecasts have proven to be systematically incorrect: core inflation remained broadly stable at 1% despite the stubbornly predicted increase, while the unemployment rate fell faster than predicted. Such forecast errors, which are also inconsistent with each other, raise serious doubts about the reliability of the ECB's current forecast of accelerating core inflation and necessitates a reflection on the inflation aim of the ECB.

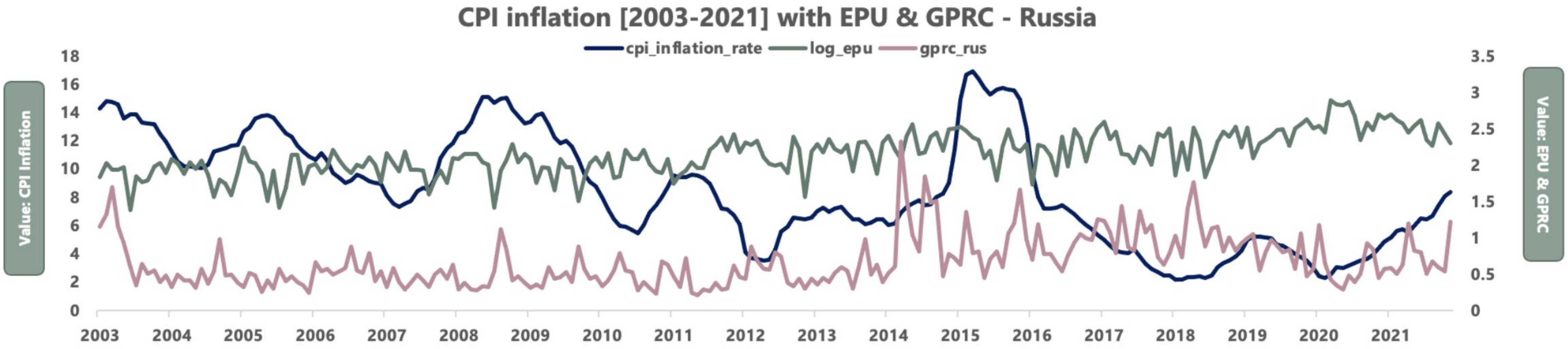
Recent projections by Eurosystem and ECB staff have substantially underestimated the surge in inflation, largely due to exceptional developments such as unprecedented energy price dynamics and supply bottlenecks. Although headline HICP inflation projections for 2020 were fairly accurate despite the emergence of the coronavirus (COVID-19) pandemic, some

CPI Inflation: BRIC Countries

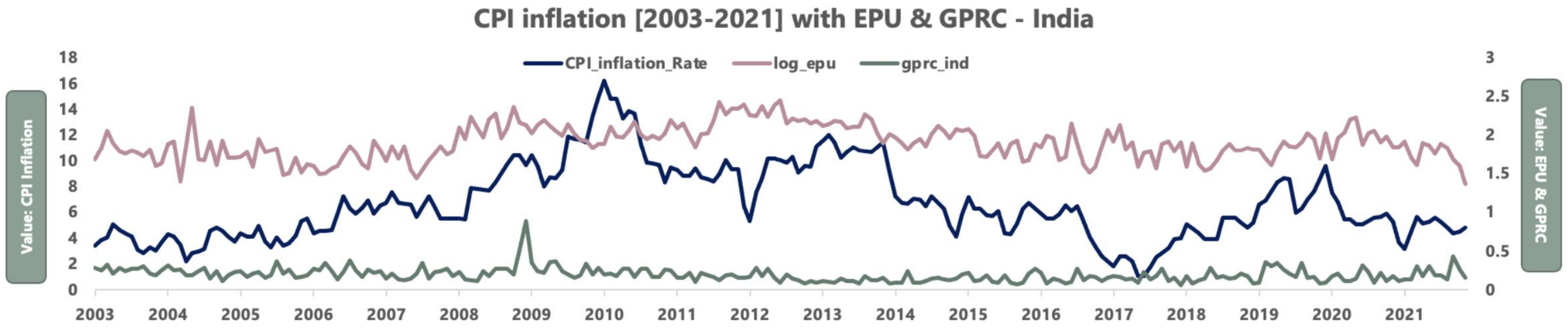


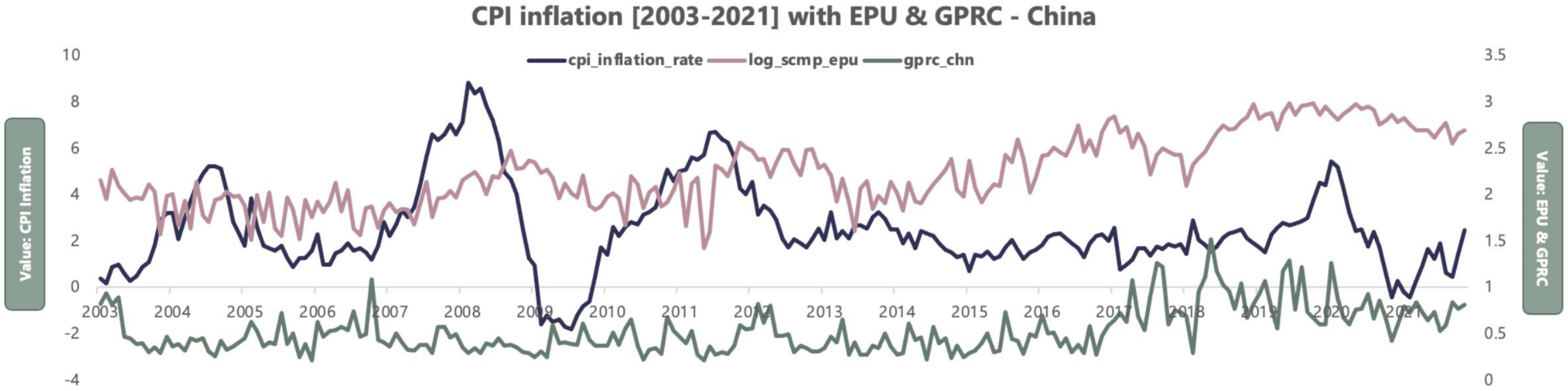
CPI Inflation – EPU & GPRC – BRIC Countries





CPI Inflation – EPU & GPRC – BRIC Countries





Global Characteristics

Countries	Series	Skewness	Kurtosis	Non-Linearity	Long-Range Dependence	Seasonality	Stationarity
Brazil	CPI Inflation	1.67	3.57	Non-linear	0.73	Non-seasonal	Non-stationary
	log(EPU)	-0.9	0.29	Non-linear	0.77	Non-seasonal	Non-stationary
	GPRC	2.06	6.60	Non-linear	0.63	Non-seasonal	Stationary
Russia	CPI Inflation	0.28	-0.98	Linear	0.78	Non-seasonal	Non-stationary
	log(EPU)	-0.03	-0.12	Non-linear	0.79	Non-seasonal	Non-stationary
	GPRC	1.41	2.46	Linear	0.79	Non-seasonal	Non-stationary
India	CPI Inflation	0.71	0.17	Linear	0.82	Non-seasonal	Non-stationary
	log(EPU)	0.12	-0.47	Linear	0.80	Non-seasonal	Non-stationary
	GPRC	2.51	15.22	Non-linear	0.73	Non-seasonal	Non-stationary
China	CPI Inflation	0.82	1.08	Non-linear	0.69	Non Seasonal	Stationary
	log(EPU)	0.23	-0.87	Non-linear	0.82	Non-seasonal	Non-stationary
	GPRC	1.28	1.51	Non-linear	0.80	Non-seasonal	Non-stationary

Causality Analysis: Wavelet Coherence Analysis

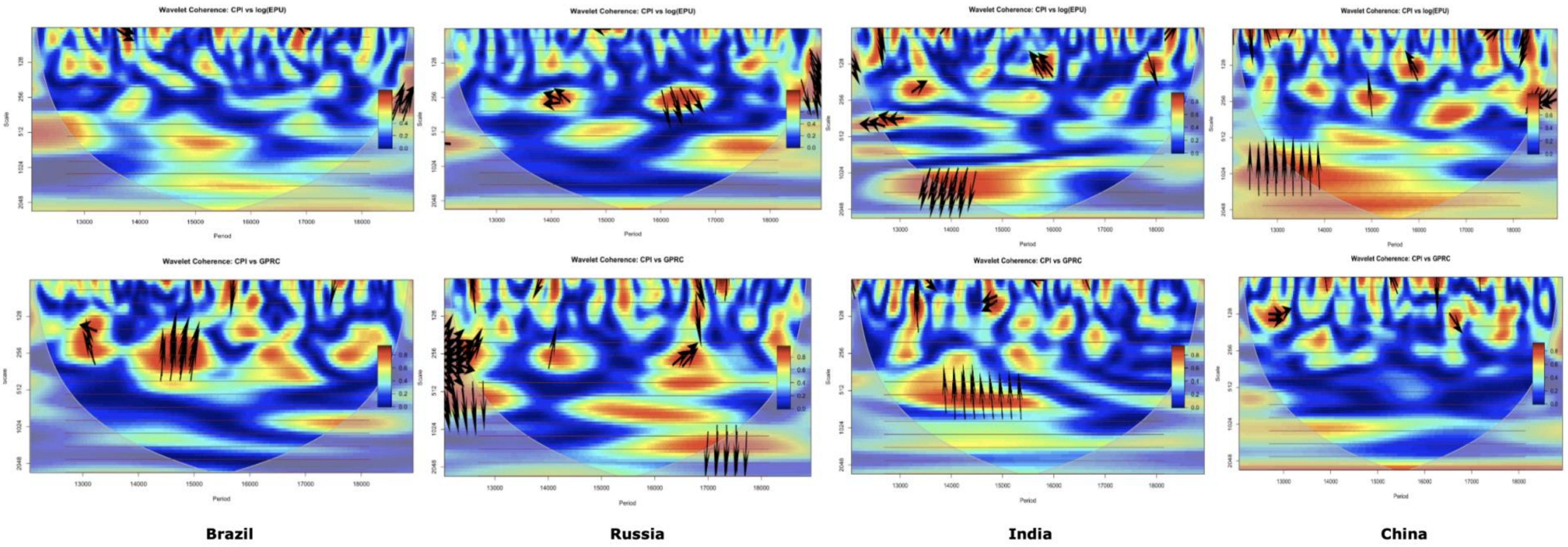


Figure 1: Wavelet coherence analysis plots - CPI inflation, log-transformed EPU (top) and CPI inflation, GPRC (bottom) for BRIC Countries

Mathematical transformations

Log transform

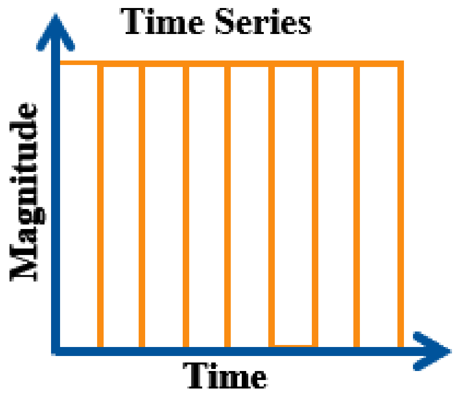
- Reduces the variability of skewed datasets.
- Highly impacted by outliers.
- Errors are symmetric on the original scale but asymmetric on the log scale.

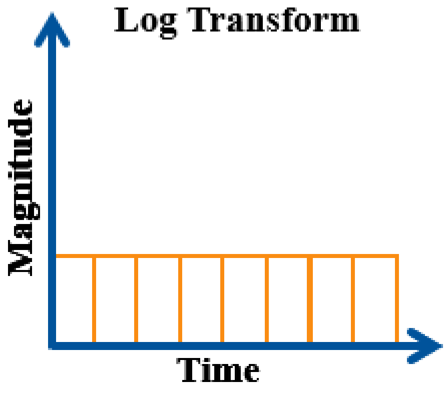
Fourier transform

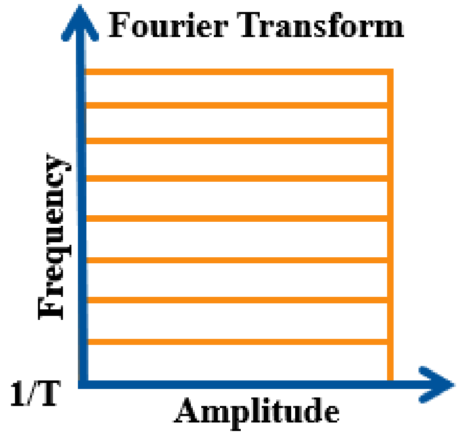
- Ideal for periodic signals.
- Represents a signal only in frequency domain
- For non-periodic signals with time-varying features, it gives averaged data, hence unsatisfactory.

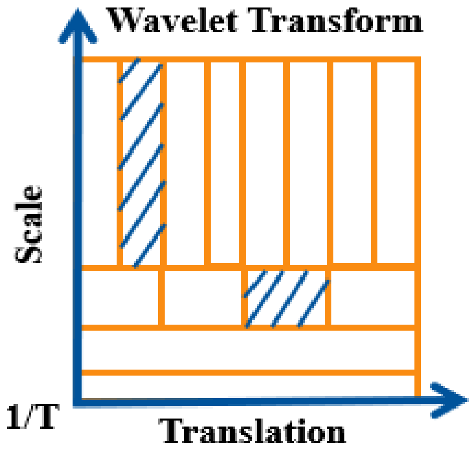
Wavelet transform

- Generalization of Fourier transform.
- It allows the independent choice of time and frequency resolution at different times and frequencies.

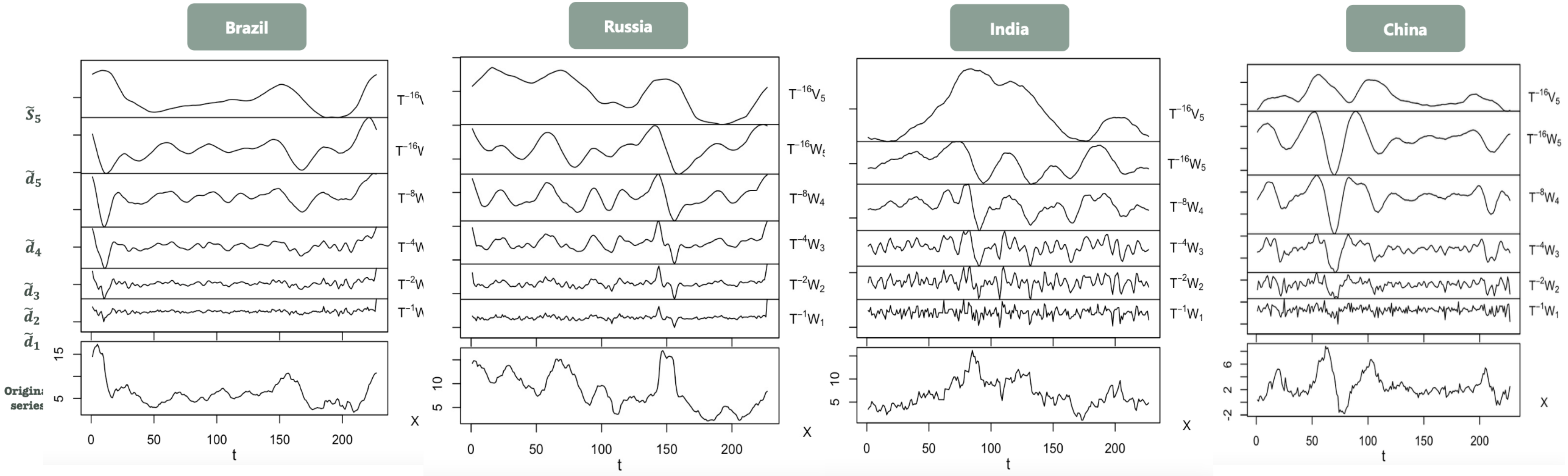








MODWT : CPI Inflation



✷The above figures show the **MODWT** (Maximal Overlap Discrete Wavelet Transform) decomposition (wavelet ($W_1, W_2, .., W_5$) and unit scale (V_1)coefficients) of the CPI Inflation series for the BRIC Countries between the period Jan-2003 and Nov-2021.It is clearly observed that at the higher scales time series is more stressed with lower frequency. $\tilde{\boldsymbol{d}}_1, \tilde{\boldsymbol{d}}_2, \ldots, \tilde{\boldsymbol{d}}_5$ represent the wavelet details and $\tilde{\boldsymbol{s}}_5$ represents the wavelet smooth..

Bandpass Filters: CPI Inflation, EPU & GPRC

- Filtering methods deal with the identification and extractions of certain features like trend and seasonalities from a time series

- Examples of different filters from macroeconomics domains:

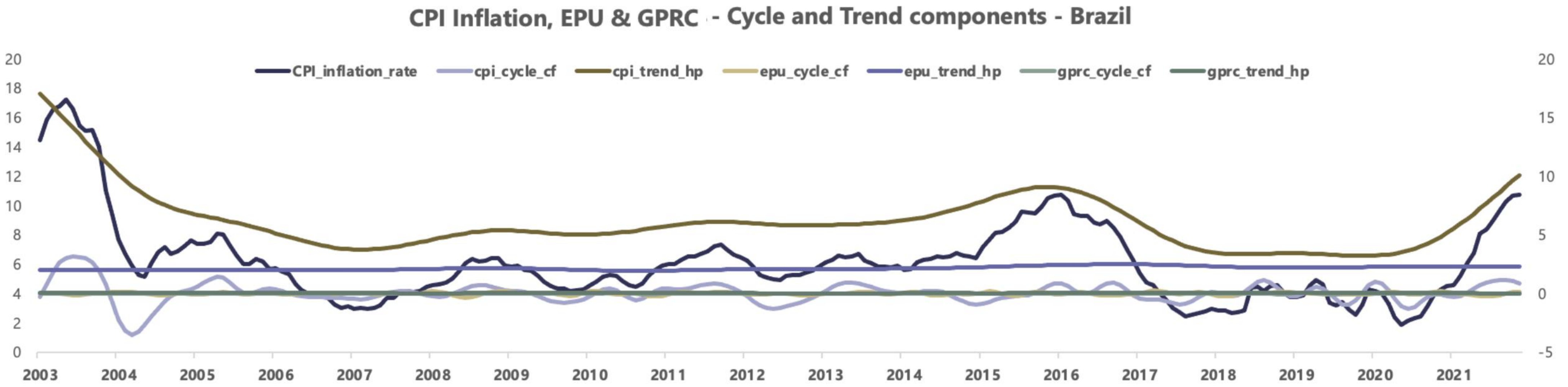
  - Hodrick – Prescott (HP) Filter
  - Christiano – Fitzgerald (CF) Filter

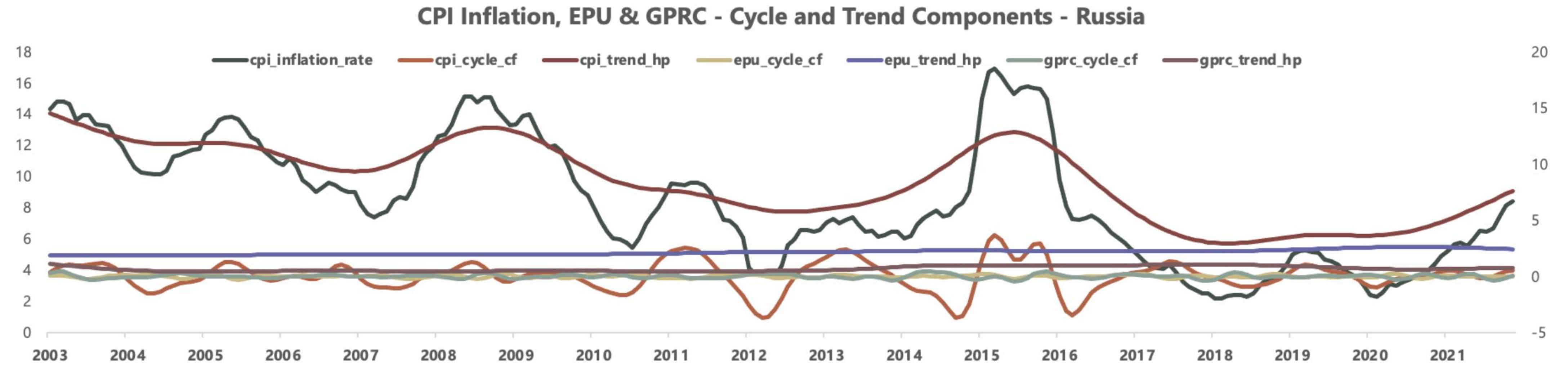
- The Hodrick and Prescott (HP) filter is widely used to de-trend macroeconomic time series (Hodrick and Prescott, 1997).

- HP filter has been explored to extract the trend component for CPI HL inflation, EPU and GPRC series.

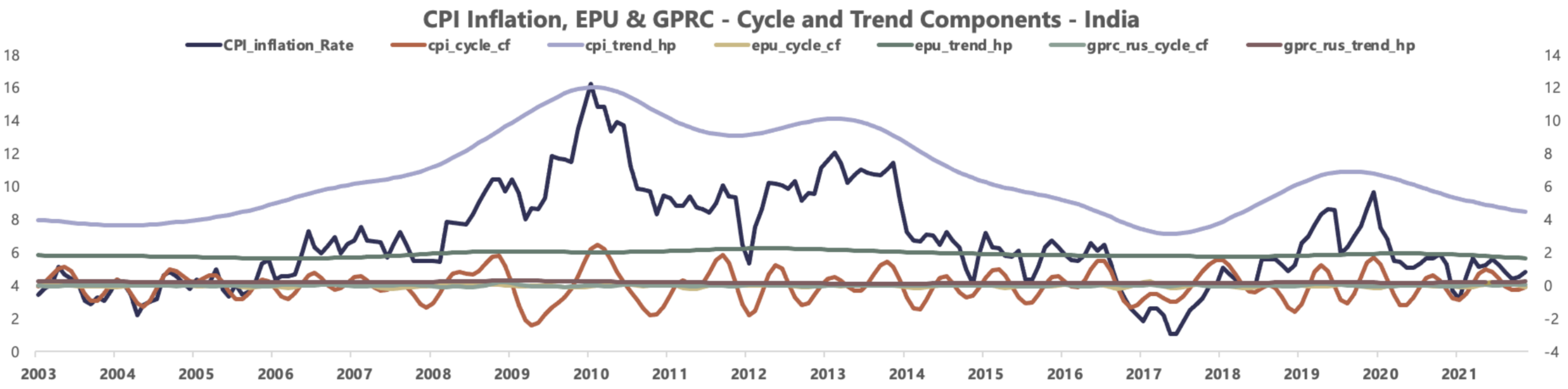
- Christiano-Fitzgerald ideal band pass filter (Christiano and Fitzgerald, 2003) has been used to extract the cyclical components from the original series.

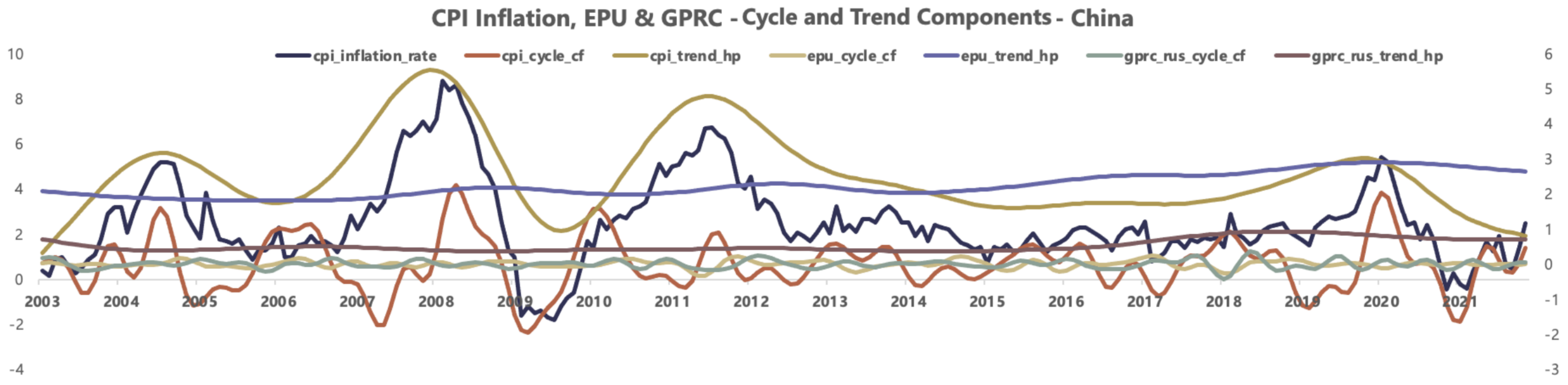
Bandpass Filters: CPI Inflation, EPU & GPRC



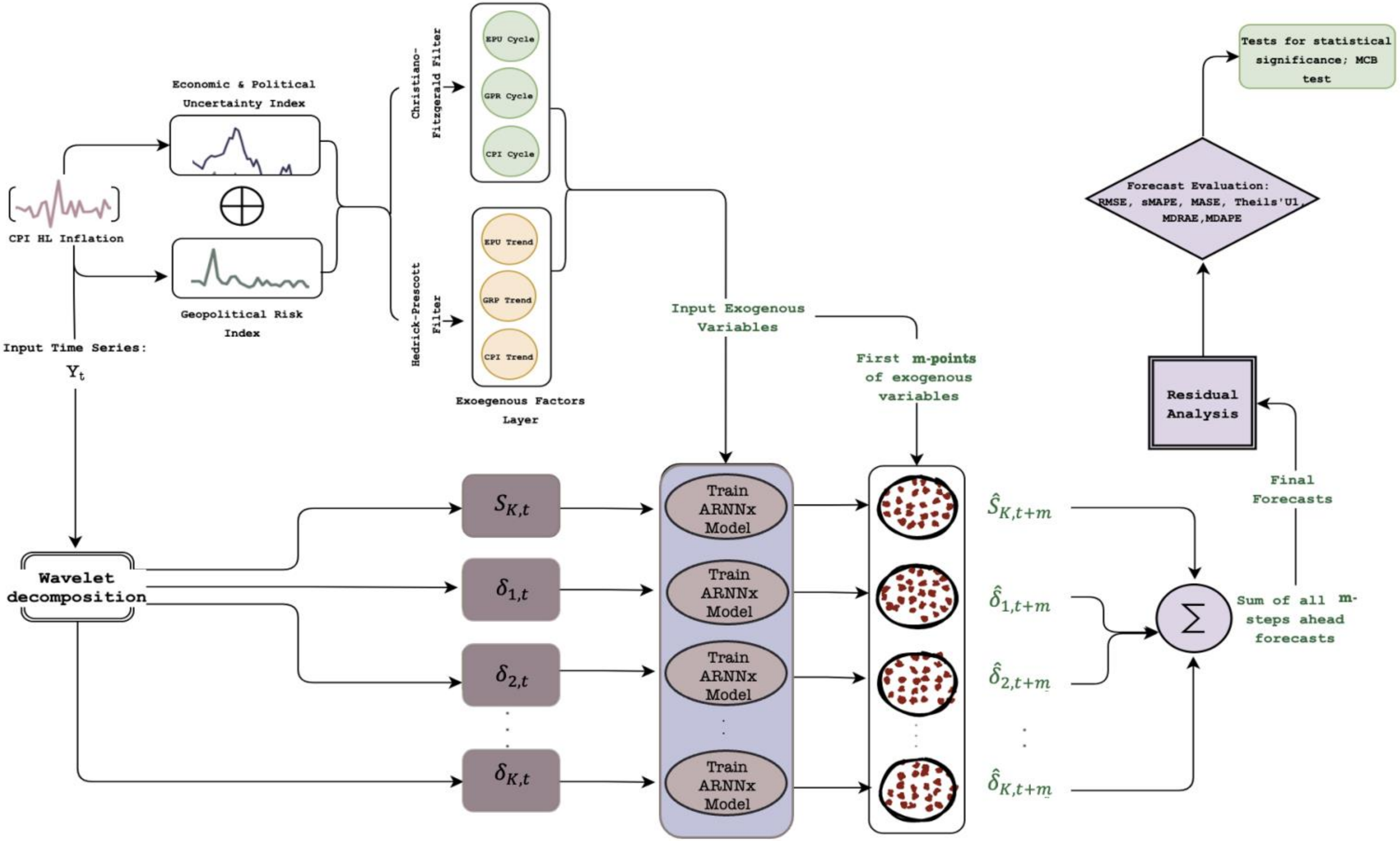


Bandpass Filters: CPI Inflation, EPU & GPRC

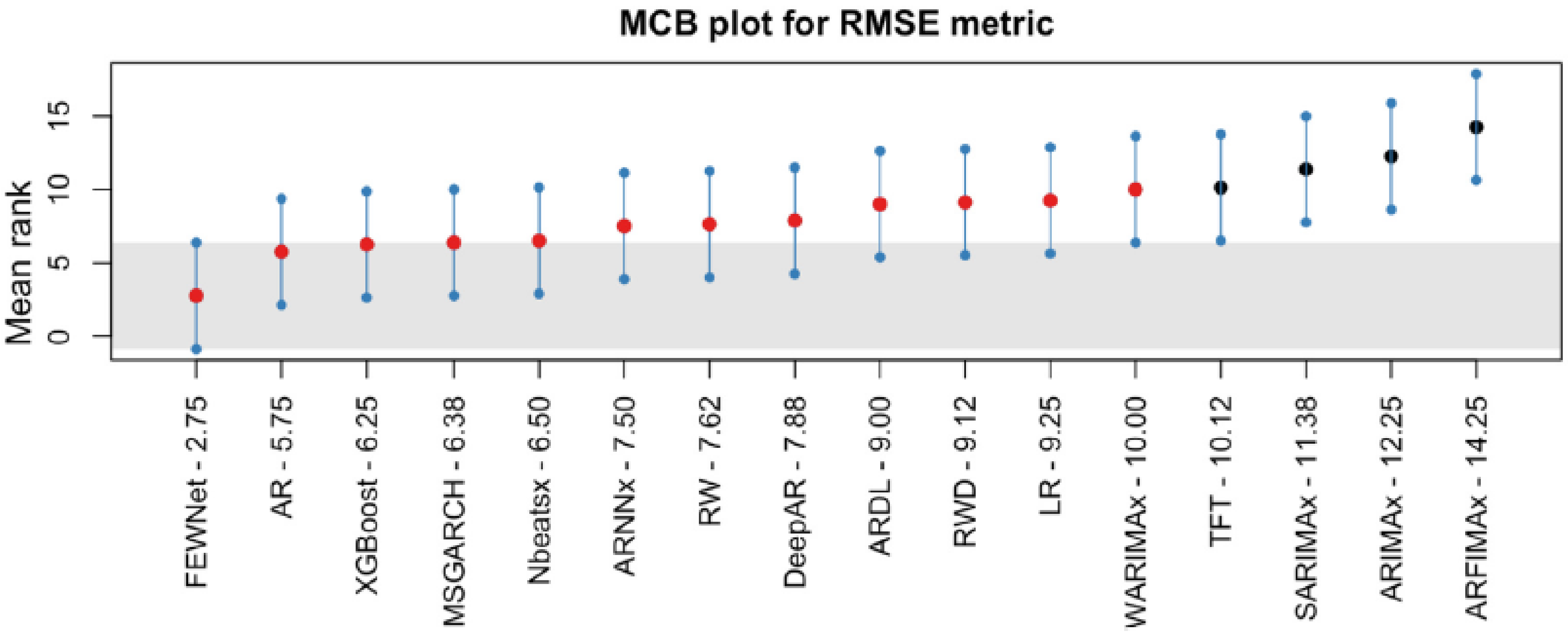




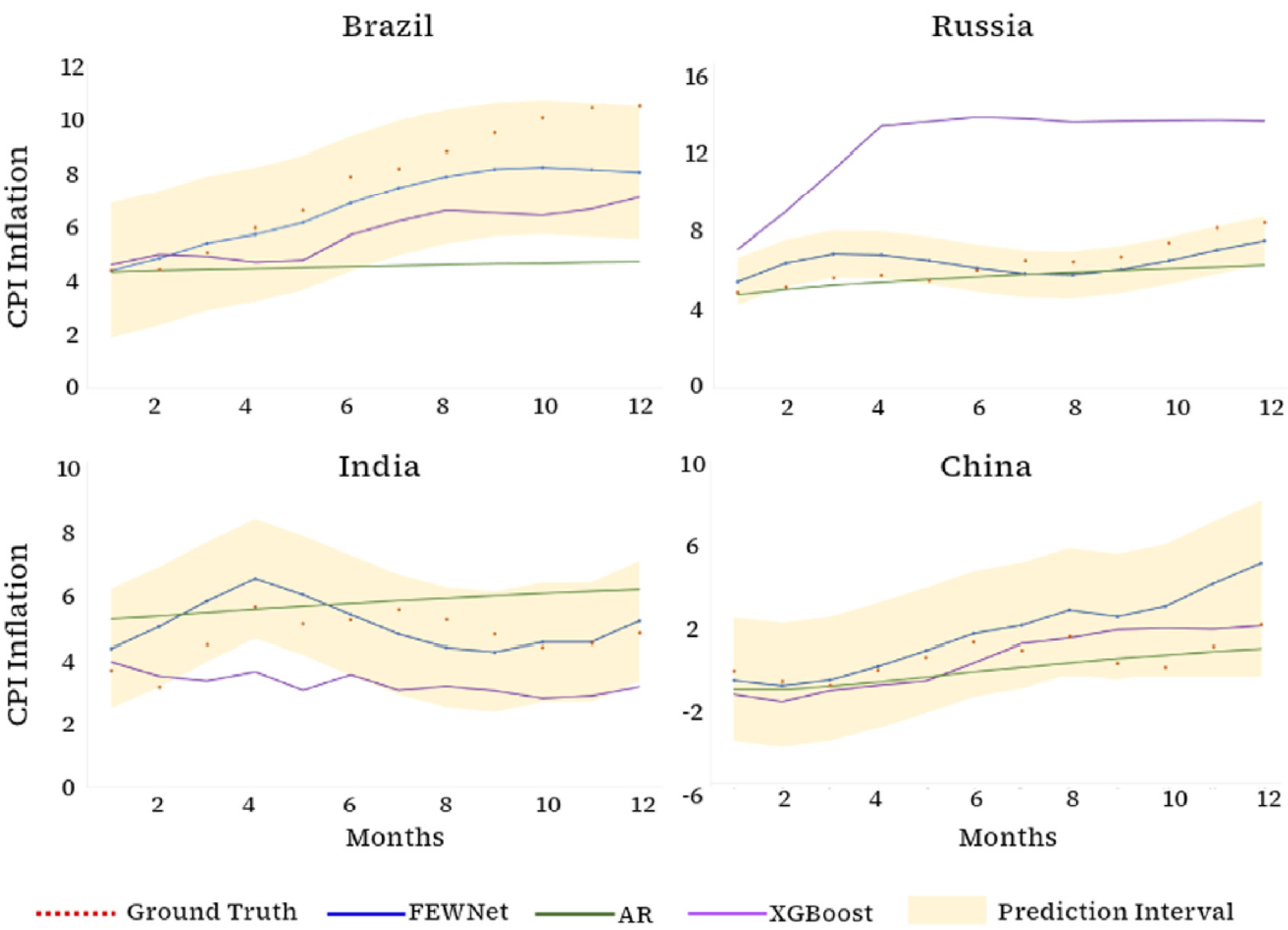
Filtered Ensemble Wavelet Neural Network (FEWNet)



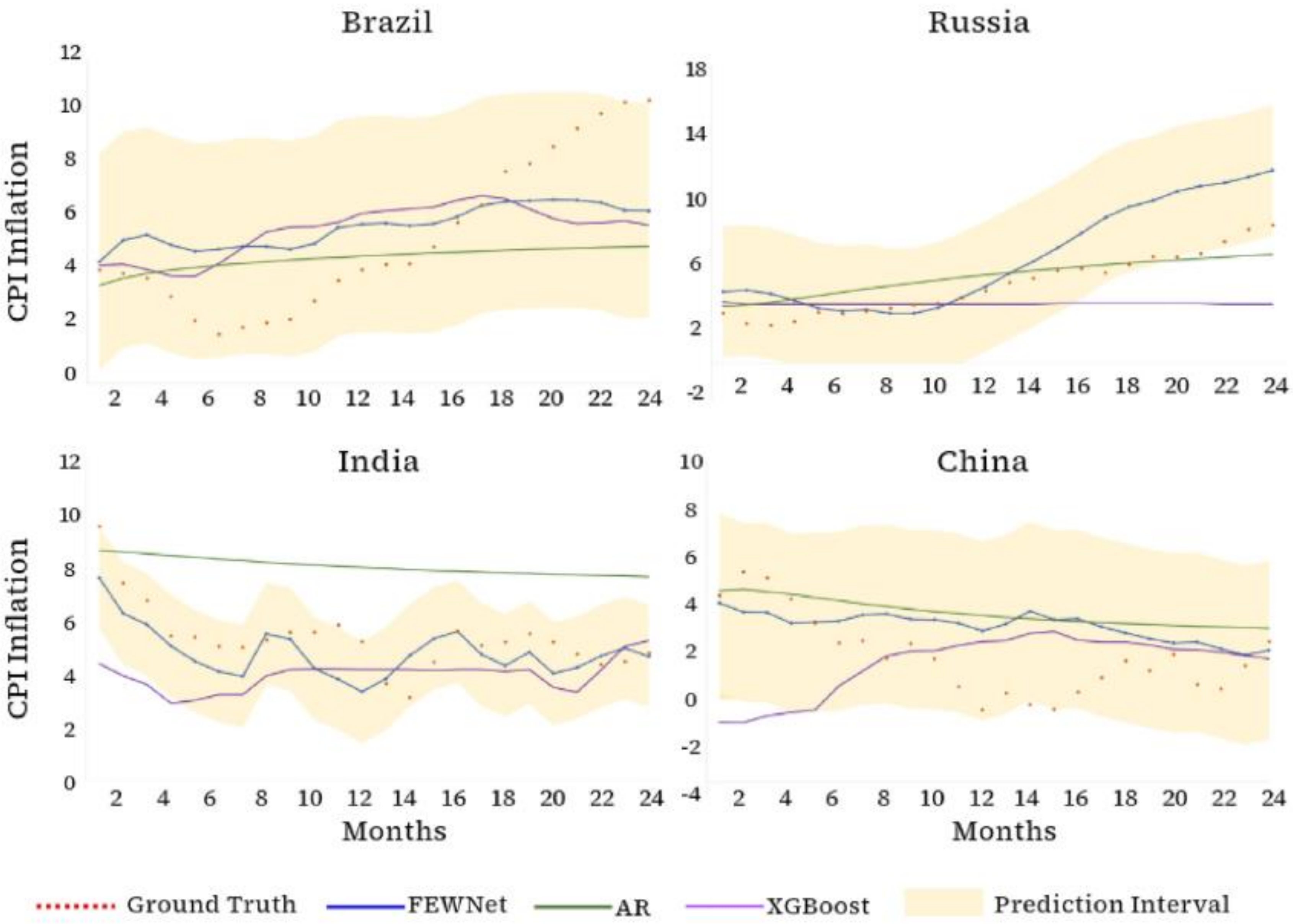
Experimental Results and Baseline Comparison



Actual vs Forecast Visualization



Actual vs Forecast Visualization



The bright future of forecasting

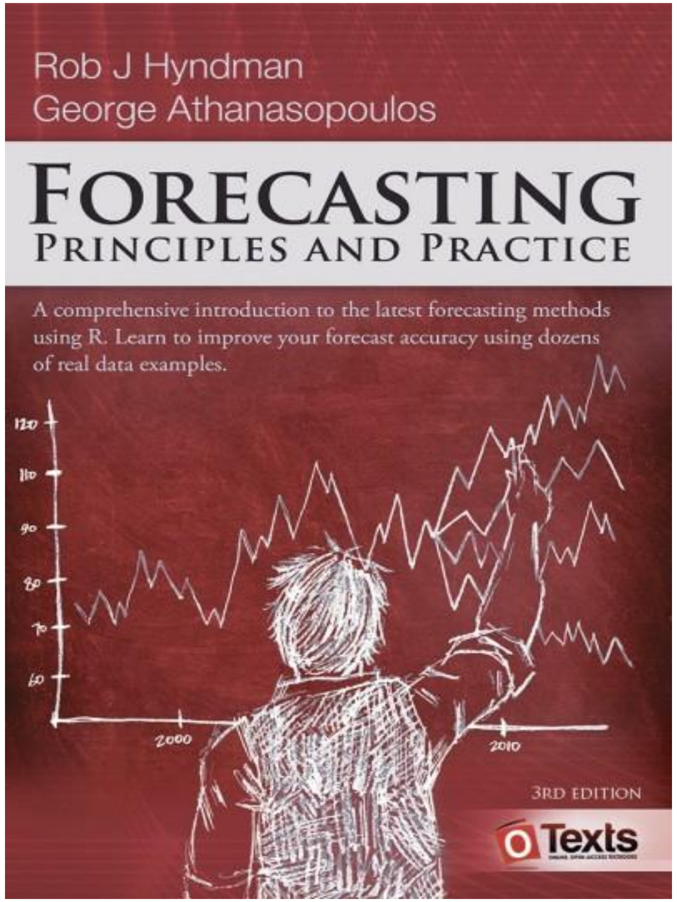
- What about interpretable neural networks and deep learning?
- Will we ever be able to forecast "black swans"?
- Does more data mean better forecasts in economics/finance?
- Will AI take over macroeconomic/financial forecasting?

Annual Review of Economics

Machine Learning Methods That Economists Should Know About

Susan Athey[1,2,3] and Guido W. Imbens[1,2,3,4]

Thank You!



Learn Forecasting from: https://otexts.com/fpp3/

Scan me for Slides