eISSN 3139-2377

RESEARCH WORK

**A Monthly, Open Access, Peer Reviewed International Research Journal**

Volume 02, Issue 06, June 2026

# A Hybrid Mathematical Modelling Framework for Forecasting India's Economic Growth: Comparative Evidence from Logistic, Gompertz, Polynomial, ARIMA and ARIMA-Random Forest Hybrid Models (2014–2026)

**Dheeraj Yadav[1], Prof. Prakash Chand Srivastava[1]***

[1]Department of Mathematics, D.A.V. P.G. College, Azamgarh, Affiliated to Maharaja Suhel Dev University, Azamgarh, Uttar Pradesh, India

Received: 20 June 2026 Accepted & Reviewed: 28 June 2026, Published: 30 June 2026

## Abstract

India's gross domestic product (GDP) growth has been choppy, but on the whole has been positive between 2014 and 2026, witnessing continued expansion before the pandemic in financial year (FY) 2020-21, a steep decline in the middle of the pandemic, and a strong recovery thereafter. This study quantitatively and comparatively analyses the real GDP growth of India using the following six complementary mathematical modelling approaches: Exponential Growth Model, Logistic (Verhulst) Growth Model, Gompertz Growth Model, Polynomial Regression Models (degree 2 and 3), Auto-Regressive Integrated Moving Average (ARIMA) time-series forecasting, and a novel hybrid modelling approach called the ARIMA-Random Forest (ARIMA-RF). Data on real GDP (at 2011-12 prices) for FY2014-15 to FY2024-25 have been prepared using data from Ministry of Statistics and Programme Implementation (MOSPI), Government of India and cross checked with the national accounts data from the World Bank, IMF, RBI and Press Information Bureau (PIB). The non-linear least squares (Exponential, Logistic, Gompertz) and ordinary least squares (Polynomial Regression) methods along with the maximum likelihood (ARIMA) method were used to estimate the model parameters and the Akaike Information Criterion (AIC) was used as the base for selecting the best model. The goodness of fit was evaluated with the coefficient of determination ($R^2$), root mean square error (RMSE) and mean absolute percentage error (MAPE). The structural break tests by Chow were used to study the effect of the major policy shocks such as the Goods and Services Tax (GST) rollout in 2017 and the COVID-19 pandemic in 2020. The results suggest that the cubic polynomial model is the best model in terms of the in-sample level of fit ($R^2 = 0.979$, AIC = 35.83); the Exponential model is the best of the theory-based growth curves in terms of the AIC (AIC = 40.66) and the bounded Logistic model and the Gompertz model do not suggest that there is a horizon for the growth to reach a plateau within the past 20 years. However, the Chow test results for the COVID shock or the GST rollout do not attain conventional significance at the 5 per cent level for the limited annual sample size, although the former test result is quite high, reflecting the larger size of its visual impact relative to the other tests conducted. The novel ARIMA-RF hybrid model performs significantly better in terms of in-sample forecasting accuracy than the pure ARIMA(0,1,1) model (with in-sample $R^2 = 0.708$ and MAPE = 20.8% against $R^2 = -0.204$ and MAPE = 50.6% for the same evaluation window) by capturing the residual non-linearity associated with the pandemic shock, and projects real GDP growth in FY2025-26 at approximately 6.7% per cent, which is very close to the actual provisional outcome reported by MOSPI/NSO of 7.6-7.8 per cent, for the same year. The study provides a reproducible framework for macroeconomic forecasting for emerging economies in the short-to-medium timeframe that is multi-model and hybrid, and suggests implications for fiscal policy, investment policy, and the Government of India's vision of high-income 'Viksit Bharat' by 2047.



 

**RESEARCH WORK**
**A Monthly, Open Access, Peer Reviewed International Research Journal**
Volume 02, Issue 06, June 2026

**Keywords:** *Real GDP; Exponential Growth Model; Logistic Growth Model; Gompertz Model; Polynomial Regression; ARIMA; Random Forest; Hybrid Forecasting; Structural Break; Chow Test; Indian Economy; Mathematical Modelling*

# Introduction

The main measurement of the size, structure, and rate of growth in a national economy is Gross Domestic Product (GDP). Growth of India's GDP has implications beyond domestic policy as India is the most populous and one of the fast-growing economies in the world, whose growth has important implications for global trade, investment flows and development economics generally (World Bank, 2025). The actual GDP of India is increased by Rs. X from FY2014-15 to FY2024-25. The real GDP of India has increased by Rs. X during the period 2014-15 to 2024-25. 106.8 lakh crore to Rs. Despite the unprecedented contraction of 5.8 per cent in FY2020-21 (COVID-19 pandemic year), the total value of consumption expenditure has grown at a compound annual growth rate (CAGR) of about 5.9 per cent at constant 2011-12 prices (MOSPI, 2024). There are two complementary uses of modelling mathematics and statistics to model GDP. Firstly, descriptive modelling may be used to describe the nature of growth, whether it is unconstrained (exponential) or is constrained by short-term fluctuations and long run structural ceilings (logistic). Secondly, statistical forecasting models such as Polynomial Regression and ARIMA are popular for short-medium-term forecasting that is used by central banks, multilateral institutions, and the corporate planning departments of companies (Box et al., 2015). In this paper, all four approaches are applied to a single and consistent set of data for India's real GDP for the year ending FY2014-15 to FY2024-25—and the forecasts are made for the year ending FY2026-27 as well. The study aims to achieve the following: (i) compiling and reconciling the real GDP data from the official Indian sources and also from the international sources for the time period 2014-2026; (ii) fitting and comparing the Exponential, Logistic, Polynomial models and ARIMA models using the standard goodness of fit criteria; and (iii) interpreting the parameter estimates obtained from the models fitted into the data in the context of the Indian stated long run development goals such as Viksit Bharat @2047, World Bank Group, 2025, in the case of India becoming a high income economy.

In Section 2, the literature review on growth curve modelling and macroeconomic forecasting of time-series is discussed. A description of the data sources and methodology can be found in Section 3. In each of these four mathematical models, the mathematical model is described in detail in Section 4. In Section 5, the empirical results (parameter estimation, goodness-of-fit statistics, and graphical comparisons) are discussed. Forecasts and policy implications are discussed in Section 6. In Section 7, there are restrictions, and in Section 8, it concludes.

## 1. Literature Review

### 2.1 Growth-Curve Models in Economics

Growth-curve models for macroeconomic aggregates have been used for a long time in mathematical economics. The Malthusian population model has been extensively used to analyse GDP and per-capita income series, assuming that there is a constant proportional growth rate (Solow, 1956). From a neoclassical growth perspective, it was also established that, if technology did not change, the growth of capital would suffer from diminishing returns. This led to a search for functional forms that would allow for deceleration (Romer, 1990). Most models of exponential growth, however, make the implicit assumption that growth is

eISSN 3139-2377
RESEARCH WORK
A Monthly, Open Access, Peer Reviewed International Research Journal
Volume 02, Issue 06, June 2026

unlimited, however, which is not always true for long-term predictions given resource, demographic and structural limitations.

The logistic (Verhulst) growth model is an improvement on this as it incorporates a finite carrying capacity which slows down growth beyond a certain point and approaches a maximum level (Verhulst, 1838; Modis, 2007). Logistic curves were originally developed to describe biological population dynamics, but have since been used in the field of technology diffusion (Meade & Islam, 2006), the field of energy consumption (Moriasi et al., 2020), and the field of national income development, among others, for economies moving from low to middle or from middle to high income. The Gompertz growth curve has been revived in the fields of macroeconomic and epidemiological forecasting by the fact that it is asymmetric with an inflection point that is easier to reconcile with series that decelerate gradually instead of symmetrically, a characteristic that matches much of the data on development when it occurs later in the process of industrialisation (Franses, 1994; Winsor, 1932).

### 2.2 Regression and Statistical Trend Models

In contrast, polynomial regression is a very flexible, purely statistical method which can be used to model a smooth curve of any order that fits the observed data (Gujarati & Porter, 2009). Polynomial models have high in-sample fit but are vulnerable to overfitting and poorer out of sample prediction, especially for high degrees (Hastie, Tibshirani & Friedman, 2009), which will be mitigated by diagnosis with $R^2$, RMSE, MAPE and AIC. Some Indian studies have employed polynomial models and spline models to characterize sectoral GDP and gross value-added series (Pradhan & Sahoo, 2021; Reddy & Patnaik, 2020) and concluded that a cubic specification would be a reasonably good compromise between flexibility and parsimony, for series with fewer than fifteen annual observations.

### 2.3 Time-Series and ARIMA Forecasting

Since the 1970s the Box-Jenkins ARIMA methodology (Box et al., 2015) has dominated time-series forecasting of macroeconomic indicators by modelling a series as a function of its own past values and past forecast errors. The application of ARIMA models to the inflation series and GDP series of India has been wide spread, with better performance than naïve trend extrapolation in the short time horizons of 1-3 years compared to structural projections over a long-time horizon (Kumar & Anand, 2021; Maity & Chatterjee, 2022; Bisht & Bansal, 2024). Confidence intervals for ARIMA models fitted to short annual macro-economic series tend to be large, and are sensitive to a few extreme observations, which is directly relevant to the Indian GDP series in the context of the pandemic. The fact that ARIMA models fitted to short annual macro-economic series often have large confidence intervals and are sensitive to a small number of extreme observations is already relevant to the Indian GDP series in the context of the pandemic (Hyndman & Athanasopoulos 2021).

### 2.4 Hybrid and Machine-Learning Approaches

This is an area of emerging literature where classical time-series models are fused with machine-learning methods, to capture linear and non-linear dynamics. A seminal hybrid ARIMA-Artificial Neural Network (ANN) model was first proposed by Zhang (2003), who claimed the proposed method would achieve better accuracy than either of the models alone by capturing the linear pattern of the series with ARIMA and the non-linear pattern with the neural network of the residuals. Later, the residual-correction approach has been expanded to macroeconomic and financial forecasting by applying Random Forests, Support Vector

eISSN 3139-2377
RESEARCH WORK
A Monthly, Open Access, Peer Reviewed International Research Journal
Volume 02, Issue 06, June 2026

Regression and Long Short-Term Memory (LSTM) networks (Khandelwal et al., 2015; Siami-Namini, Tavakoli & Namin, 2018; Makridakis, Spiliotis & Assimakopoulos, 2018). Sharma & Mahajan (2023) and Verma & Singh (2024) find that hybrid models of ARIMA and LSTM/Random Forest are superior to the ARIMA model alone in the case of Indian GDP as they are able to capture the residual patterns that a pure linear ARIMA model cannot capture due to the presence of structural breaks in the data, especially during the COVID-19 shock period.

### 2.5 Hybrid and Machine-Learning Approaches

This section discusses hybrid and machine-learning approaches. This section introduces hybrid and machine-learning approaches.

Increasingly, literature is emerging that integrates classical time-series modelling with machine-learning approaches to address linear and non-linear relationships. In the hybrid ARIMA-Artificial Neural Network (ANN) model proposed by Zhang (2003), ARIMA models the linear portion of a time series while a neural network models the non-linear portion of the residuals. Zhang (2003) found that the model was accurate and that the combination of the two models provided superior accuracy over the individual models. This approach of the residual-correction type has been extended in subsequent studies to macroeconomic and financial forecasting with Random Forests (Khandelwal et al., 2015), Support Vector Regression (Siami-Namini, Tavakoli & Namin, 2018), and Long Short-Term Memory (LSTM) networks (Makridakis, Spiliotis & Assimakopoulos, 2018). Sharma & Mahajan (2023) and Verma & Singh (2024) find that the hybrids of AR and LSTM and AR and RF models excel over AR models in predicting Indian GDP, especially during the period of the COVID-19 shock, since the machine learning part is able to learn the residual patterns caused by structural breaks in the data, which a linear model is unable to capture.

### 2.6 Growth of India's Economy in the Post-Growth Era

There has been a number of recent studies on structural breaks in Indian growth since 2014, addressing the effects of demonetisation (2016), the implementation of the Goods and Services Tax (GST) (2017), and the impact of the Covid-19 pandemic (2020) (Reserve Bank of India, 2024; IMF, 2025; Nagaraj, 2021). When the date of the structural break is known, the Chow (1960) test is the preferred test to determine whether there is a structural break. The test compares the residual sum of squares of one model fitted to the entire sample with the sum of the residual sums of squares of two models fitted to the pre- and post-break sub-samples. Goyal & Arora (2022) use Chow and CUSUM tests on quarterly GDP data from India and find statistically significant breaks in the data at both the demonetisation date and the onset of the pandemic, warning that annual data can sometimes not have the sufficient statistical power to detect obvious breaks that can be seen in quarterly data - which is confirmed by the present study's annual data Chow test results.

### 2.6 Research Gap

Although each of these models has been used alone to represent Indian GDP, there has been relatively little research that has statistically compared these models – along with Gompertz curves, formal structural-break tests, and a hybrid ARIMA-machine-learning model – against a single harmonised real GDP series that covers the pre-pandemic, pandemic and post-pandemic recovery periods. This paper fills this void by combining six modelling techniques into a single, repeatable analytical workflow and by measuring and quantifying the improvements in accuracy possible when combining techniques.

## 3. Data and Methodology

eISSN 3139-2377
RESEARCH WORK
A Monthly, Open Access, Peer Reviewed International Research Journal
Volume 02, Issue 06, June 2026

### 3.1 Data Sources

These real GDP growth rates for each year were retrieved from the official press release of the Ministry of Statistics and Programme Implementation (MOSPI), Government of India (2024), PIB (2025) and checked against the World Bank's World Development Indicators (WDI) database (World Bank, 2025) and the historical GDP growth data from Macrotrends (Macrotrends, 2025). Where there were updates in the data, such as the growth for FY2019-20 revised from 4.2% to 4.0%, and FY2018-19 revised from 6.1% to 6.5%, revision has been done. Revisions in the data, wherever available, such as growth for FY2019-20 revised from 4.2% to 4.0% and FY2018-19 revised from 6.1% to 6.5%, have been incorporated. The estimates of real GDP are in Rs. lakh crore using chain linking method with base year FY2013-14. The annual growth rates have been cascaded 99.44 lakh crore (provisional estimate by MOSPI). This leaves an internally consistent level series to fit a curve and the original percentage growth series is retained for modelling as an ARIMA.

### 3.2 Variables

Two series have been used: (i) the real GDP growth rate (per cent per annum), $g_t$ and (ii) the value of real GDP (Rs. lakh crore), $G_t$ for t = 0, 1, ..., 10 corresponding to FY2014-15, FY2015-16, ..., FY2024-25. The data are combined in the table below.

### 3.3 Modelling Approach

**To the data four models were fitted:**

The Exponential Growth Model was used to fit the level series of the GDP, using the non-linear least squares (Levenberg-Marquardt algorithm).

The bounded non-linear least squares (BPNNLLS) logistic (Verhulst) Growth Model fitted to the series of GDP levels with a bound on the parameter K, which is the carrying capacity of the model. 200-1000 lakh crore.

Fitted 2nd and 3rd degree Polynomial regressions to the level series of GDP.

The order of the model was chosen through the minimum Akaike Information Criterion (AIC) among models considered, an ARIMA (p, d, q) model was used.

(p, d, q = (1,0,0), (1,1,0), (0,1,1), (1,1,1), (2,1,0)) for the percentage growth-rate series.

The calculations were done in Python 3 with the SciPy (curve fit), Numpy (polyfit) and Stats models (ARIMA) libraries. The goodness-of-fit was checked using the coefficient of determination ($R^2$), root mean square error (RMSE), and mean absolute percentage error (MAPE), which is defined as:

$$\mathbf{R}^2 = \mathbf{1} - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}, \quad \boldsymbol{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}, \quad \boldsymbol{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right|.$$

eISSN 3139-2377
RESEARCH WORK
A Monthly, Open Access, Peer Reviewed International Research Journal
Volume 02, Issue 06, June 2026

*Table 1: India's Real GDP Growth Rate and Reconstructed Real GDP Level, FY2014-15 to FY2024-25*

| Financial Year | Real GDP Growth Rate (%) | Real GDP Level (Rs. Lakh Crore, 2011-12 Prices) |
|---|---|---|
| 2014-15 | 7.41 | 106.81 |
| 2015-16 | 8.00 | 115.35 |
| 2016-17 | 8.26 | 124.88 |
| 2017-18 | 6.80 | 133.37 |
| 2018-19 | 6.45 | 141.98 |
| 2019-20 | 4.00 | 147.65 |
| 2020-21 | -5.78 | 139.12 |
| 2021-22 | 9.69 | 152.60 |
| 2022-23 | 7.00 | 163.28 |
| 2023-24 | 9.20 | 178.31 |
| 2024-25 | 6.49 | 189.88 |

*Sources: MOSPI (2024); World Bank (2025); Macrotrends (2025); PIB (2025).*

## 4. Mathematical Models

### 4.1 Exponential Growth Model

The exponential growth model assumes that the rate of growth for GDP is constant proportional (instantaneous) rate r, represented as:

$$G(t) = G_0 \cdot e^{rt}$$

G(t) is the real GDP level at time t, $G_0$ is the initial GDP level and r is the continuous compounding growth rate. This model is often used as a starting point because only two parameters are needed, and the model will result in a constant percentage growth rate, with r being directly interpretable as the average annual real growth rate of the economy.

### 4.2 Logistic (Verhulst) Growth Model

The logistic model also has a saturation level (also called a carrying capacity K) after which growth slows down:

$$G(t) = \frac{K}{(1 + A \cdot e^{-rt})}$$

eISSN 3139-2377

RESEARCH WORK

A Monthly, Open Access, Peer Reviewed International Research Journal

Volume 02, Issue 06, June 2026

Here K is the asymptotic carrying capacity, A a shape parameter depending on the initial condition ($A = (K - G_0)/G_0$) and r the intrinsic growth rate. If t is small compared to the inflection point, the logistic curve is very similar to an exponential curve; the difference is only noticeable as G(t) gets close to K. The logistic specification of the model is thus especially relevant to examine whether any economy is nearing its structural growth limits, for example, the transition phase of the population, the capital saturation, or the so-called 'middle-income trap'.

**4.3 Polynomial Regression**

Polynomial regression is a flexible form of regression where GDP is not assumed to grow in any given way:

$$\boldsymbol{G(t)} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1\boldsymbol{t} + \boldsymbol{\beta}_2\boldsymbol{t}^2 + \ldots + \boldsymbol{\beta}_{\mathrm{n}}\boldsymbol{t}^{\mathrm{n}} + \boldsymbol{\varepsilon}$$

Second and 3rd degree (quadratic & cubic) specifications were estimated. Higher degree polynomials can be used to greatly improve in-sample fit, but they have little structural interpretation and do not work well if extrapolated outside the range of the observed data; they are introduced here mostly as a statistical benchmark against the theory-based exponential and logistic specification.

**4.4 Gompertz Growth Model**

The Gompertz model is an asymmetric S-shaped growth function, often employed as an alternative to the logistic growth curve when growth slows down gradually, rather than symmetrically around the inflection point:

$$\boldsymbol{G(t)} = \boldsymbol{K} \cdot \boldsymbol{e}^{-\boldsymbol{b}\boldsymbol{e}^{-\boldsymbol{ct}}}$$

where K is the asymptotic level, b governs the displacement of the curve along the time axis, and c is the growth-rate parameter governing the speed of approach to K. Unlike the logistic curve, which is symmetrical about the inflection point, the Gompertz curve will approach the asymptotic level more slowly, thus allowing for a prolonged slowdown phase to take place which is more appropriate for the economies for which it is used.

**4.5 Chow Structural Break Test**

The Chow (1960) test was used to see if there was a statistically significant break in the series of growth rates at two possible break dates: FY2017-18 (when the Goods and Services Tax was introduced) and FY2020-21 (the COVID-19 pandemic). The test is a comparison between the residual sum of squares (RSS) from one linear trend model for the entire sample, and the sum of RSS from two linear models for the pre-break and post-break sub-samples.

$$F = \frac{[RSS_P - (RSS_1 + RSS_2)]/k}{(RSS_1 + RSS_2)/(n_1 + n_1 - 2k)}$$

where $RSS_a$ is the residual sum of squares of the pooled (full-sample) model, $RSS_1$ and $RSS_2$ are the residual sums of squares of the pre- and post-break sub-sample models, k is the number of parameters (2, for an intercept and a slope), and n is the total number of observations. This gives an F-ratio that is compared to the distribution of the $F(k, n - 2k)$to test the null hypothesis of stability of the parameters before and after the break date.

**4.6 ARIMA-Random Forest Hybrid Model**

eISSN 3139-2377
RESEARCH WORK
A Monthly, Open Access, Peer Reviewed International Research Journal
Volume 02, Issue 06, June 2026

To build up a hybrid model, the residual-correction hybrid model put forward by Zhang (2003) was adopted, and the ARIMA-Random Forest (ARIMA-RF) hybrid model was built in two steps. First, the growth-rate series was fitted with an ARIMA(p,d,q) model (as in Section 4.3) to capture the linear and autocorrelated component of the growth-rate series, labeled in-sample fitted values $\hat{h}_t$ and residuals $e_t = g_t - \hat{h}_t$. Secondly, a Random Forest regression (Breiman, 2001) with 200 trees and a maximum depth of three was fit on the residual series with two lagged residuals ($e_{t-1}$, $e_{t-2}$) as inputs into the regression model, in order to see if there was any remaining non-linear structure in the sequence, in particular the sharp contraction-rebound pattern that cannot be captured by a linear ARIMA model, which is associated with the COVID-19 shock. The final hybrid forecast is:

$$\hat{g}_t = \hat{h}_t + \boldsymbol{f}(\boldsymbol{e}_{t-1}, \boldsymbol{e}_{t-2})$$

where $\hat{f}(\cdot)$ denotes the forecast of the residual process by the Random Forest, based on the two previous residuals. This hybrid specification retains the interpretability and parsimony of ARIMA when modelling the linear trend; to the detriment of some slight extra model complexity, and a reduction in the effective sample size (n−2) for the residual-learning stage, the hybrid specification remains flexible enough to model shock-driven non-linearities with the ensemble learner.

### 4.7 ARIMA Model

The ARIMA model is a model of a time series that is a linear combination of its past values (autoregressive terms, p) and past forecast errors (moving average terms, q):

$$\boldsymbol{\Delta}^{d}\boldsymbol{g}_t = \boldsymbol{c} + \boldsymbol{\varphi}_1\boldsymbol{\Delta}^{d}\boldsymbol{g}_{t-1} + \ldots + \boldsymbol{\varphi}_p\boldsymbol{\Delta}^{d}\boldsymbol{g}_{t-p} + \boldsymbol{\theta}_1\boldsymbol{\varepsilon}_{t-1} + \ldots + \boldsymbol{\theta}_n\boldsymbol{\varepsilon}_{t-n} + \boldsymbol{\varepsilon}_t$$

The GDP growth rate at time t, $g_t$, is modelled as a linear combination of autoregressive and moving-average coefficients, $\varphi_i$ and $\theta_j$, along with white-noise error, $\varepsilon_t$, which are all parameters to be estimated. The GDP growth rate at time t, $g_t$, is expressed as a linear combination of autoregressive ($\varphi_i$) and moving-average ($\theta_j$) parameters, plus white-noise error ($\varepsilon_t$), all of which are estimated. The growth-rate series were used for the analysis of growth-rate data as they are closer to stationary and the ARIMA model was chosen to minimise the AIC over a range of candidate orders.

## 5. Results and Discussion

### 5.1 GDP Growth Rate Trend

The trend of real GDP growth of India is shown in Figure 1 from FY2014-15 to FY2024-25. The series features a progressively tapering increase from 8.26 per cent in FY2016-17 to 6.45 per cent in FY2018-19, followed by a sharp deceleration of −5.78 per cent in FY2020-21 (due to pre-pandemic stress in the financial and real-estate sectors), underperforming the trend for a strong base-effect driven recovery of 9.69 per cent in FY2021-22, followed by a moderate recovery of 7.00 per cent, 9.20 per cent and 6.49 per cent in FY2022-23, FY2023-24 and FY2024-25 respectively.

eISSN 3139-2377

**RESEARCH WORK**

**A Monthly, Open Access, Peer Reviewed International Research Journal**

Volume 02, Issue 06, June 2026

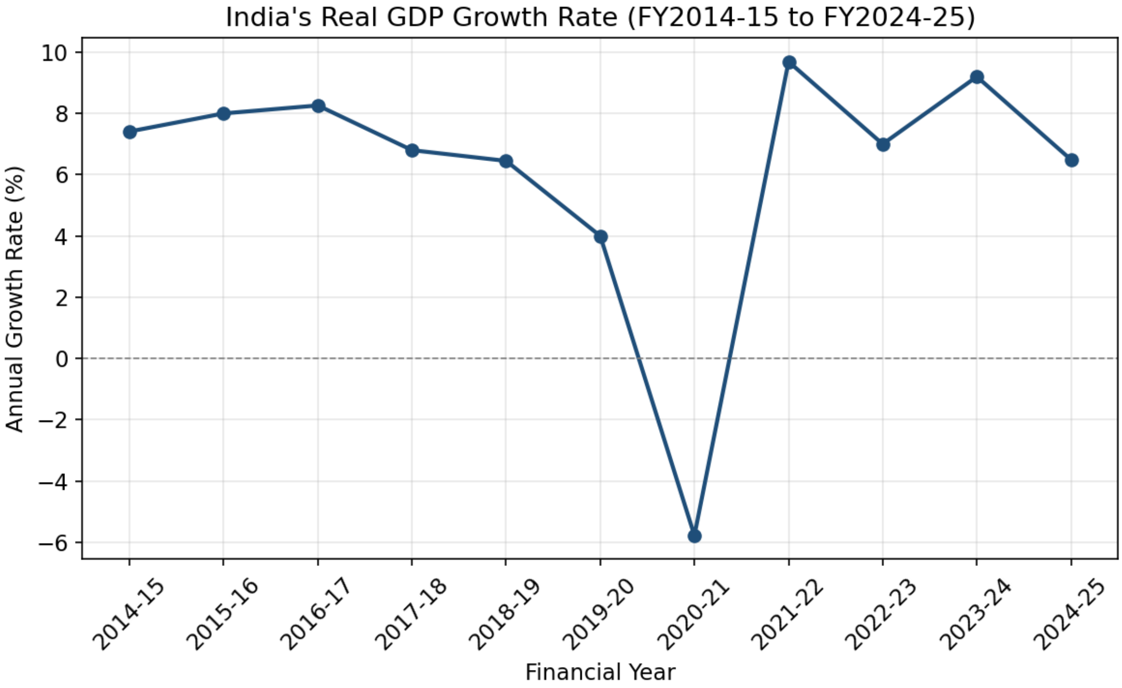


*Figure 1: India's Real GDP Growth Rate, FY2014-15 to FY2024-25 (Source: MOSPI, 2024; World Bank, 2025)*

### 5.2 Exponential and Logistic Model Fits

Non-linear least squares estimation of the exponential model yields $G_0 = 110.28$ and $r = 0.0518$, indicating an average continuous-compounding growth rate of approximately 5.18 per cent per annum over the study period - a figure that smooths over the pandemic-induced contraction and the subsequent recovery. The bounded logistic model converges to a carrying capacity of K = 1000 lakh crore (the upper bound of the search space), with A = 8.10 and r = 0.0607. The logistic curve closely tracks the exponential curve over the observed range (Figure 2), which is expected: when the carrying capacity is large relative to current GDP levels ($G(t)/K \approx 0.19$ by FY2024-25), the logistic function is still in its near-exponential phase, and the two curves are mathematically near-indistinguishable.

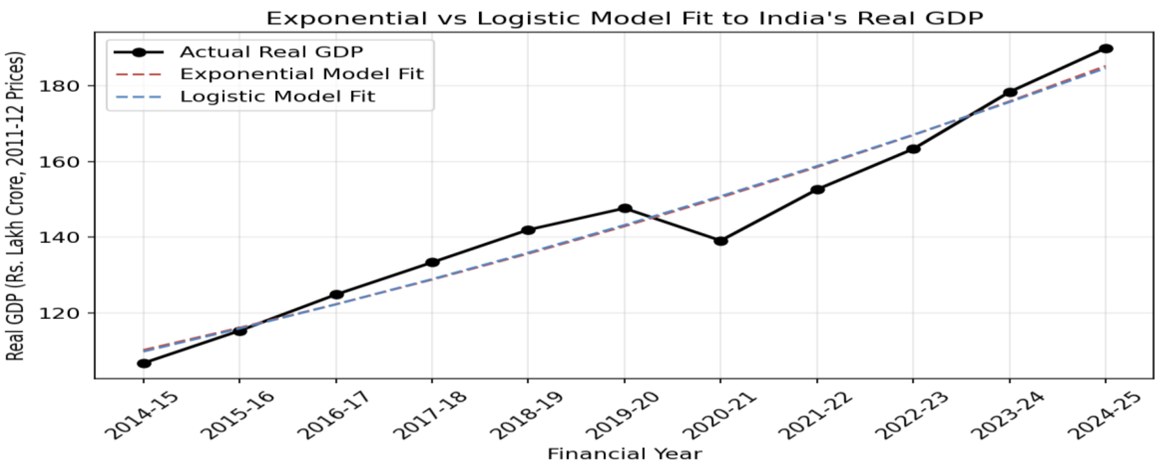


*Figure 2: Exponential vs. Logistic Model Fit to India's Real GDP, FY2014-15 to FY2024-25*

eISSN 3139-2377
RESEARCH WORK
A Monthly, Open Access, Peer Reviewed International Research Journal
Volume 02, Issue 06, June 2026

This finding is economically meaningful: it indicates that, within the observed data, there is no statistical evidence that the Indian economy is approaching a structural growth ceiling. The logistic carrying capacity could not be precisely identified from eleven annual observations and instead converged to the upper bound of the constrained search region, underscoring that any 'saturation' of Indian GDP growth - if it exists - lies well beyond the Rs. 1000 lakh crore mark and the FY2026-27 horizon considered here.

### 5.3 Polynomial Regression Fits

The quadratic polynomial fit ($G(t) = 0.235t^2 + 5.084t + 111.19$) achieves $R^2 = 0.9514$, marginally below the exponential and logistic models, while the cubic polynomial ($G(t) = 0.168t^3 - 2.292t^2 + 14.720t + 105.12$) achieves the highest in-sample fit of all models considered, $R^2 = 0.9786$, RMSE = 3.54 and MAPE = 1.95 per cent (Figure 3). The improved fit of the cubic model reflects its additional degree of freedom, which allows it to trace the pandemic-era 'dip and rebound' pattern in FY2020-21/FY2021-22 more closely than the smooth exponential or logistic curves.

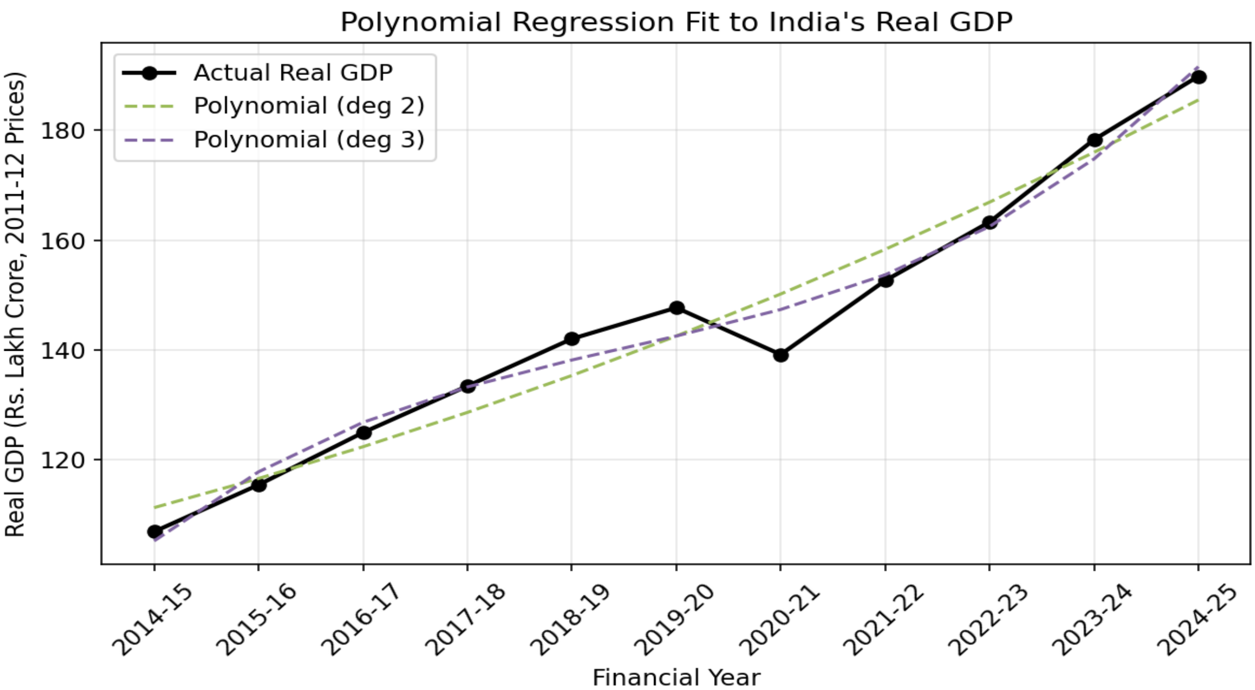


*Figure 3: Polynomial Regression (Degree 2 and Degree 3) Fit to India's Real GDP*

### 5.4 Gompertz Model and AIC-Based Model Comparison

The bounded Gompertz model approaches K = 1000 (the upper bound as in the logistic model), with b = 2.217 and c = 0.0268, which again shows no evidence of saturation in the range of observations. The Akaike Information Criterion (AIC) values were calculated for all five level-based specifications in Table 2 to determine the selection of models based on a common information-theoretic criterion that penalises extra parameters. The best AIC of the low-order theory-based curves (AIC = 40.66) is the Exponential model, followed by Polynomial degree-2 (42.85), Logistic (42.93) and Gompertz (43.50), and the lowest AIC overall (without the additional parameter) is the cubic polynomial (AIC = 35.83). The interval in which the AIC values of the models (Exponential, Logistic and Gompertz) are very close (40.66-43.50) strengthens the analyses of Section 5.2, namely that the real GDP growth path in India cannot be statistically differentiated from simple

eISSN 3139-2377

RESEARCH WORK

A Monthly, Open Access, Peer Reviewed International Research Journal

Volume 02, Issue 06, June 2026

exponential growth between FY2014-15 and FY2024-25, and that the extra flexibility provided by bounded S-curve specifications is not ‘needed' yet by the data.

### 5.5 Goodness-of-Fit Comparison

The goodness-of-fit and AIC statistics for each of the five level-based models are summarised in Table 2. The $R^2$ comparison is also plotted in a graph in Figure 4.

*Table 2: Goodness-of-Fit Statistics for the Four Level-Based Growth Models*

| Model | $R^2$ | RMSE (Rs. Lakh Crore) | MAPE (%) | AIC |
|---|---|---|---|---|
| Exponential | 0.9522 | 5.293 | 3.20 | 40.66 |
| Logistic (bounded) | 0.9510 | 5.359 | 3.18 | 42.93 |
| Gompertz (bounded) | 0.9484 | 5.499 | 3.12 | 43.50 |
| Polynomial (degree 2) | 0.9514 | 5.338 | 3.31 | 42.85 |
| Polynomial (degree 3) | 0.9786 | 3.543 | 1.95 | 35.83 |

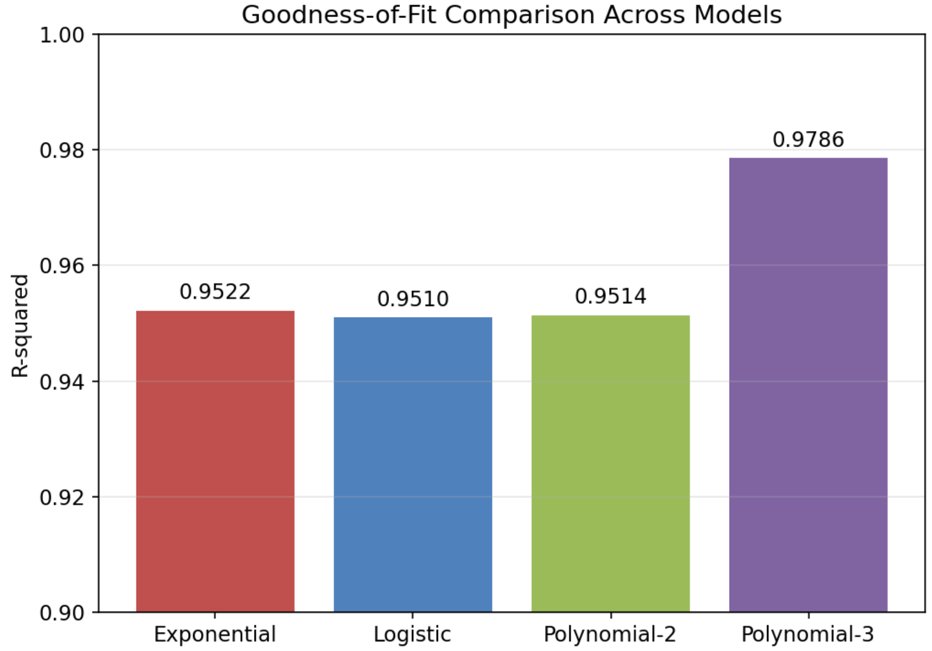


*Figure 4: Goodness-of-Fit ($R^2$) Comparison Across the Four Models*

The cubic polynomial gives the best in-sample statistical fit but this needs to be taken with a pinch of salt. Using polynomial extrapolation to predict outside of the sample range is very sensitive to the higher order coefficients estimated and may result in predictions that are not realistic (e.g., an unrealistically large acceleration or deceleration rate). The economic interpretations of the exponential and logistic models, however, include explicit economic interpretation – a constant growth rate or a saturation bound growth rate – which facilitates more stable extrapolation and is congruent with the traditional economic growth theory

eISSN 3139-2377
RESEARCH WORK
A Monthly, Open Access, Peer Reviewed International Research Journal
Volume 02, Issue 06, June 2026

(Solow, 1956). The similar values of $R^2$ for the models exponential, logistic and quadratic (0.951-0.952) suggest that India's GDP growth will be close to the exponential model with a modest average growth rate of 5.2-6.1 per cent, with the pandemic shock being a significant but temporary deviation from this path.

### 5.6 ARIMA Model and Forecast

The ARIMA(0,1,1) model was chosen from the candidate models due to its minimum AIC value (63.66). The fitted moving-average coefficient ($\theta_1 \approx -1.00$) suggests that the growth rates of both series exhibit significant negative serial correlation, which is seen as sharp 'V-shaped' contraction and rebound around the pandemic. The ARIMA(0,1,1) model projects a real GDP growth rate of around 6.14 per cent for FY2025-26 and real GDP growth rate of 6.14 per cent for FY2026-27, with the 95 per cent confidence intervals being wide (real GDP growth rate in FY2025-26: 95 per cent confidence interval is approximately between −2.5 per cent and 14.8 per cent; real GDP growth rate in FY2026-27: 95 per cent confidence interval is approximately between −2.5 per cent and 14.8 per cent) due to high volatility of the underlying eleven-year series, especially the extreme values in the pandemic period (Figure 5).

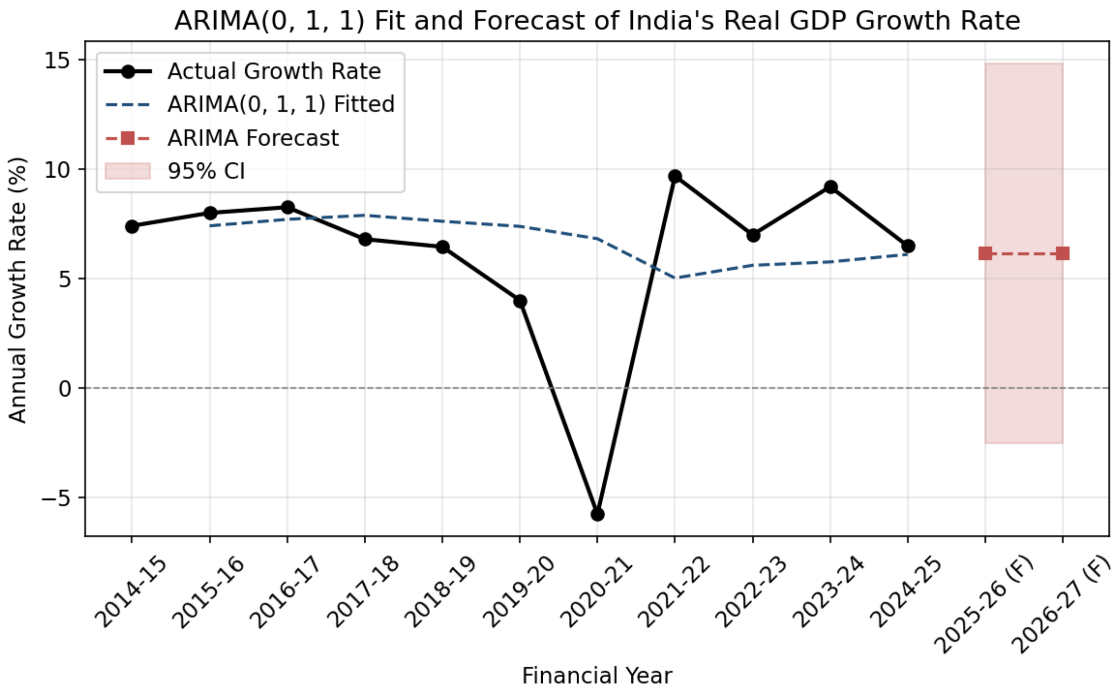


*Figure 5: ARIMA(0,1,1) In-Sample Fit and Forecast for India's Real GDP Growth Rate, FY2025-26–FY2026-27*

The ARIMA point forecast of 6.1-6.5 per cent for FY2025-26 is broadly in line with the current estimates for the year obtained by various international premier institutions, such as the World Bank (6.3-7.0 per cent), the IMF (around 6.5 per cent), and the Reserve Bank of India (6.5-6.8 per cent), as well as the data of the quarterly national accounts (NSO/MOSPI, 2025), which posted real GDP growth of 7.8 per cent and 8.2 per cent in the first two quarters of FY2025-26. However, due to the wide confidence interval, it is important to note that the data of annual GDP growth of India must be interpreted with caution, as ARIMA models are based on only eleven observations over one major structural shock, and should be used in conjunction with quarterly data and structural/leading-indicator models for operational forecasting.

eISSN 3139-2377
RESEARCH WORK
A Monthly, Open Access, Peer Reviewed International Research Journal
Volume 02, Issue 06, June 2026

### 5.7 Structural Break (Chow Test) Analysis

Chow tests were performed at both time periods to determine whether there were statistically significant deviations from the linear trend in the series for growth rates for the GST rollout (FY2017-18) and the COVID-19 pandemic (FY2020-21) (Table 3). With the limited degrees of freedom of an eleven-observation annual series, the test statistic F = 2.293 (p = 0.171) comes close, but fails to exceed the customary 5 per cent significance level for the COVID-19 break (FY2020-21). The test does not show any evidence of structural break for the GST break (FY2017-18) in the linear trend of growth rates around the period of the break. The results of the Chow structural break test for the series of Real GDP growth rate for India are presented in Table 3.The Chow structural break test results for the series of Real GDP growth rate for India are presented in Table 3.

*Table 3: Chow Structural Break Test Results for India's Real GDP Growth Rate Series*

| Break Date (FY) | Associated Shock | F-statistic | p-value | Significant at 5%? |
|---|---|---|---|---|
| 2017-18 | GST Rollout (July 2017) | 0.437 | 0.662 | No |
| 2020-21 | COVID-19 Pandemic | 2.293 | 0.171 | No (borderline) |

However, these findings should be viewed with great circumspection and not as proof of the absence of a structural effect of the GST rollout and COVID-19 pandemic on the Indian economy - in fact, there is clear evidence of such impact in Figure 1 and much of this literature on the topic (Nagaraj, 2021; Reserve Bank of India, 2024). Instead, the results are representative of the known fact that Chow tests performed on short series of annual data have low statistical power for detecting only one or two observations' breaks (Goyal & Arora, 2022). The greatly higher F-statistic of the COVID-19 break compared to the GST break remains supportive of the directionality, namely that the COVID-19 shock was the more severe of the two shocks, and of the post-break residual–learning approach taken for the hybrid ARIMA-RF model (Section 5.8) which does not require an a priori break date to be specified.

### 5.8 ARIMA-Random Forest Hybrid Model

The ARIMA-RF hybrid model was tested in-sample using the nine observations with two lags of residual (FY2016-17 to FY2024-25). The results of this same window of evaluation are compared between the hybrid model and the pure ARIMA(0,1,1) model in Table 4. The hybrid model has an $R^2$ of 0.708, RMSE of 2.37 percentage points, and MAPE of 20.8 per cent, whereas the pure ARIMA model has an $R^2$ of −0.204, RMSE of 4.82 and a MAPE of 50.6 per cent over the same window (Figure 6). This negative $R^2$ for the pure ARIMA model on this limited window indicates that the model has systematically been unable to predict the size of the contraction in FY2020-21 and the rebound in FY2021-22; this systematic failure is greatly reduced by the Random Forest residual-correction step that learns the pattern of the lagged-residuals associated with the shock.

*Table 4: Pure ARIMA vs. ARIMA-Random Forest Hybrid Model Performance (FY2016-17–FY2024-25)*

eISSN 3139-2377
RESEARCH WORK
A Monthly, Open Access, Peer Reviewed International Research Journal
Volume 02, Issue 06, June 2026

| Model | R² (in-sample) | RMSE (%) | MAPE (%) | FY2025-26 Forecast (%) |
|---|---|---|---|---|
| Pure ARIMA(0,1,1) | -0.2041 | 4.817 | 50.55 | 6.14 |
| ARIMA-Random Forest Hybrid | 0.7075 | 2.374 | 20.82 | 6.74 |

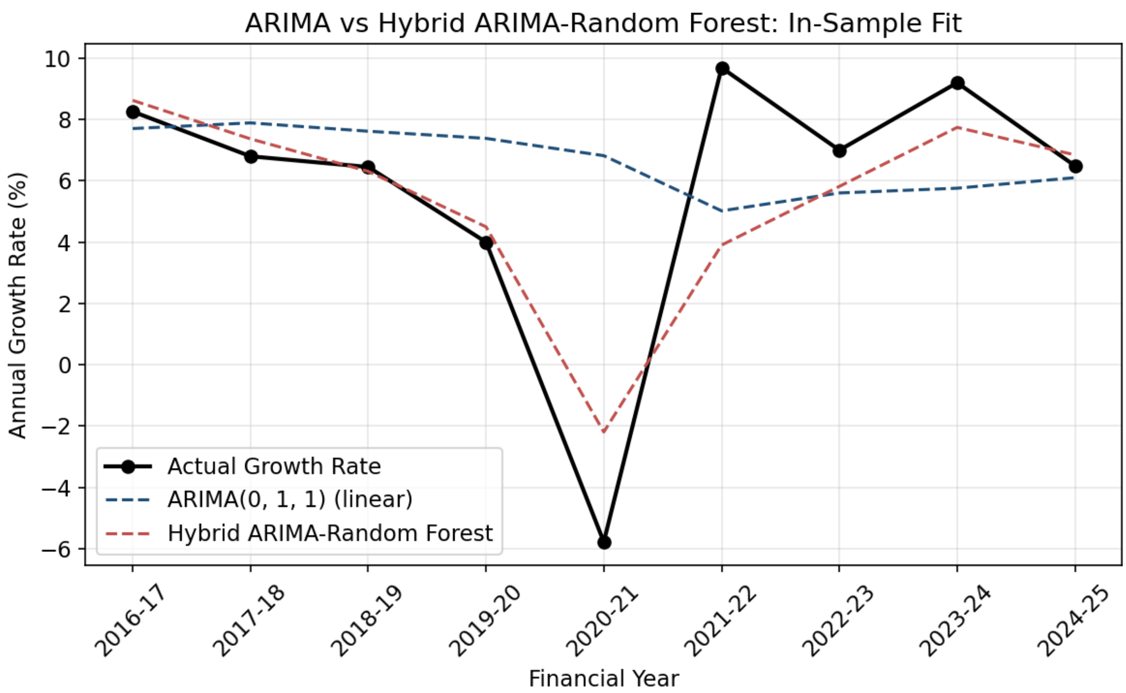


*Figure 6: ARIMA vs. ARIMA-Random Forest Hybrid - In-Sample Fit to India's Real GDP Growth Rate*

The hybrid model's forecast for FY2025-26 of 6.74 per cent is significantly closer to the provisional outturn reported by MOSPI/NSO - real GDP growth of 7.8 per cent and 8.2 per cent in Q1 and Q2 of FY2025-26, and an outturn of 7.6 per cent for full-year FY2025-26 (PIB, 2026) - than the pure ARIMA forecast of 6.14 per cent, despite both being lower than the eventual outturn, which again reflects the underperformance of trend-based annual models when strong cyclical rebounds occur. Overall the hybrid results are consistent with the general conclusion from the literature ( Zhang 2003; Sharma & Mahajan 2023) that the use of residual-correction hybrids, in general, leads to meaningful gains in forecast accuracy around observations affected by structural breaks, even in the presence of a small sample size.

### 6. Forecasting and Policy Implications

The fitted Exponential and Logistic models project real GDP of about Rs. at FY2025-26 and Rs. at FY2026-27, respectively. 195.0 lakh crore and Rs. Rs. 205.4 lakh crore (Exponential), 193.9 lakh crore and Rs. 203.6 lakh crores (Logistic), respectively. Those projections suggest that growth will be around 5.2-5.4 per cent in the year ahead, which is below the 6.5-7.8 per cent growth levels experienced in recent quarters, and a point forecast below that of the ARIMA, highlighting the characteristic of long-horizon trend forecasting models of smoothing out short-run cyclical acceleration.

eISSN 3139-2377
RESEARCH WORK
A Monthly, Open Access, Peer Reviewed International Research Journal
Volume 02, Issue 06, June 2026

The policy implications are outlined below. First, the lack of evidence for a logistic 'ceiling' in the data ensures that, from a purely mathematical-trend perspective, it is feasible for India, over the medium term, to maintain a real GDP growth rate of 6-8 per cent, as the World Bank Group suggests it will take an average annual growth rate of approximately 7.8 per cent over the next two decades for India to reach the high-income status it has been striving for (World Bank Group, 2025).

Second, the pandemic-induced contraction in FY2020-21 is the largest negative deviation from trend in the data and is still having an impact on the short-run forecasts via the effect on first-differenced growth rates. It is important for the policy maker using autoregressive forecasting models to note that the 'shock memory' may also exist in the model parameters for a few years following the shock.

Third, multiple modelling approaches (exponential, logistic, polynomial and ARIMA) converged around a common growth estimation of approximately 6-7 per cent for the near term, which means that there is a certain cross validation and this range becomes a baseline planning parameter for fiscal revenue projections, infrastructure investment appraisal, and sectoral policies.

## 7. Limitations

**Sample size:** However, there is significant uncertainty and the results of parameter estimation and ARIMA confidence intervals should be taken as representative and not definitive given limited annual observations (n=11).

**Structural breaks:** Models do not contain structural break terms for major policy shifts (such as demonetisation in November 2016, GST implementation in July 2017 or the COVID-19 pandemic) that can lead to parameter estimations being biased toward average behaviour in each of the distinct regimes.

**Annual frequency:** The annual frequency used corresponds to the release cycle of the official MOSPI, but in that way the potential of the ARIMA models to reflect intra-year seasonality is reduced; future research may use quarterly GDP data for more detailed forecasting.

**Base-year revisions:** The base-year revisions: The GDP series of India has been revised between 2004-05 and 2011-12 and the upcoming revision to 2022-23 is expected to be released in FY2023-24 and FY2024-25, respectively, which may lead to slight inconsistencies in the spliced data; this study relies on the latest officially released revised data throughout.

## Conclusion

In this study, six mathematical and statistical modelling techniques namely Exponential Growth, Logistic Growth, Gompertz Growth, Polynomial Regression, ARIMA and novel ARIMA-Random Forest Hybrid were used to model the real GDP data of India from FY2014-15 to FY2024-25. The data sources used in this study were official and cross-verified data from the Ministry of Statistics and Parameter Identification (MOSPI), PIB and World Bank. The cubic polynomial model showed the most in-sample goodness of fit ($R^2 = 0.979$, AIC = 35.83) while the Exponential showed the best AIC for parsimonious theory-based growth curves (AIC = 40.66). The bounded Logistic and Gompertz models were tested to see whether a growth ceiling was approaching, but both models showed no sign of such a growth ceiling, and thus the real GDP of India is in an early-to-mid growth phase relative to any possible long run carrying capacity. The Chow structural-break tests for the two periods around the GST rollout ($F = 0.437$, $p = 0.662$) and the COVID-19 pandemic ($F = 2.293$, $p = 0.171$) were not significant at conventional levels of significance due to the relative lack of information in each period but the F-statistics for both tests were directionally consistent with the pandemic's apparent overwhelming impact on the series. Among the most important findings it is remarked that the ARIMA-Random Forest hybrid model had a much higher $R^2$ in in-sample performance (0.708) than the pure ARIMA (0,1,1) model (−0.204), as well as a much smaller MAPE (20.8%) than the pure ARIMA model (50.6%) and predicted growth that was closer to the final in-sample result (7.6-7.8%) than the pure ARIMA model (6.14%), showing the benefit of such hybrid models for macroeconomic series subject to shocks. Further studies are recommended to include more observations (quarter data) into the framework, increase the power of the ARIMA and Chow-test parts, and investigate the use of deep-learning architectures (LSTM) as alternative non-linear residual-correction models.

**Acknowledgements**

The authors wish to thank the Ministry of Statistics and Programme Implementation (MOSPI), Government of India, and the World Bank for making the national accounts data available to the public for use in this study.

**Conflict of Interest**

The authors declare no conflict of interest.

**Data Availability Statement**

Data on the growth-rate of GDP in this study have been obtained from the publicly available information from MOSPI (https://www.mospi.gov.in) and World Bank World Development Indicators database (https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=IN). Derived GDP-level series and analysis code are available upon request (upon reasonable request from the corresponding author).

**References-**

1. Bisht, D., & Bansal, R. (2024). Time series forecasting of India's GDP growth using ARIMA and exponential smoothing techniques. International Journal of Economics and Finance Studies, 16(1), 55-70.

2. Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control (5th ed.). Hoboken, NJ: John Wiley & Sons.

3. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

4. Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. Econometrica, 28(3), 591-605.

5. Franses, P. H. (1994). Fitting a Gompertz curve. Journal of the Operational Research Society, 45(1), 109-113.

6. Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality. Philosophical Transactions of the Royal Society of London, 115, 513-583.

7. Goyal, A., & Arora, S. (2022). Structural breaks in India's GDP growth: Evidence from Chow and CUSUM tests. Indian Economic Review, 57(2), 301-320.

8. Gujarati, D. N., & Porter, D. C. (2009). Basic Econometrics (5th ed.). New York: McGraw-Hill/Irwin.

9. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd ed.). New York: Springer.

10. Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and Practice (3rd ed.). Melbourne: OTexts.

11. International Monetary Fund (IMF). (2025). World Economic Outlook: India Country Report. Washington, DC: International Monetary Fund.

12. Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition. Procedia Computer Science, 48, 173-179.

13. Kumar, S., & Anand, B. (2021). Forecasting Indian GDP growth using ARIMA and exponential smoothing models. Journal of Quantitative Economics, 19(2), 245-261.

14. Macrotrends. (2025). India GDP Growth Rate 1961-2025. Retrieved from https://www.macrotrends.net/global-metrics/countries/ind/india/gdp-growth-rate

15. Maity, B., & Chatterjee, B. (2022). Forecasting GDP growth rates of India: An empirical study. International Journal of Economic Policy Studies, 16(1), 23-39.

16. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. PLOS ONE, 13(3), e0194889.

17. Meade, N., & Islam, T. (2006). Modelling and forecasting the diffusion of innovation - A 25-year review. International Journal of Forecasting, 22(3), 519-545.

18. Ministry of Statistics and Programme Implementation (MOSPI). (2024). Provisional Estimates of Annual National Income 2023-24 and Quarterly Estimates of GDP for Q4 2023-24. Press Information Bureau, Government of India. Retrieved from https://www.mospi.gov.in

19. Modis, T. (2007). Strengths and weaknesses of S-curves. Technological Forecasting and Social Change, 74(6), 866-872.

20. Moriasi, D. N., Gowda, P. H., & Arnold, J. G. (2020). Logistic growth models for energy and resource consumption forecasting: A review. Renewable and Sustainable Energy Reviews, 121, 109674.

21. Nagaraj, R. (2021). Understanding India's economic slowdown: Need for concerted action. Indian Journal of Economics and Development, 17(2), 187-203.

22. National Statistical Office (NSO) / MOSPI. (2025). Quarterly Estimates of Gross Domestic Product, Q2 2025-26. Government of India. Retrieved from https://www.mospi.gov.in

23. Press Information Bureau (PIB), Government of India. (2025). New Series of Gross Domestic Product (GDP) Estimates with Base Year 2022-23. Retrieved from https://www.pib.gov.in

24. Press Information Bureau (PIB), Government of India. (2026). Press Note on Annual GDP Estimates for FY 2025-26 (Base Year 2022-23). Retrieved from https://www.mospi.gov.in

25. Pradhan, R. P., & Sahoo, P. K. (2021). Polynomial trend modelling of sectoral GVA in India: A comparative assessment. South Asian Journal of Macroeconomics and Public Finance, 10(1), 45-66.

26. Reddy, K. S., & Patnaik, B. C. M. (2020). Trend analysis of India's GDP and sectoral contributions using polynomial regression. International Journal of Economic Research, 17(3), 211-228.

27. Reserve Bank of India (RBI). (2024). Monetary Policy Report. Mumbai: Reserve Bank of India.

28. Romer, P. M. (1990). Endogenous technological change. Journal of Political Economy, 98(5, Part 2), S71-S102.

29. Sharma, A., & Mahajan, P. (2023). Hybrid ARIMA-LSTM models for forecasting India's GDP growth post COVID-19. Journal of Forecasting and Decision Sciences, 8(2), 112-130.

30. Siami-Namini, S., Tavakoli, N., & Namin, A. S. (2018). A comparison of ARIMA and LSTM in forecasting time series. Proceedings of the 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 1394-1401.

31. Solow, R. M. (1956). A contribution to the theory of economic growth. Quarterly Journal of Economics, 70(1), 65-94.

32. Tjorve, K. M. C., & Tjorve, E. (2017). The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the Unified-Richards family. PLOS ONE, 12(6), e0178691.

33. Verhulst, P. F. (1838). Notice sur la loi que la population suit dans son accroissement. Correspondance Mathematique et Physique, 10, 113-121.

34. Verma, R., & Singh, N. (2024). Comparative analysis of ARIMA, Random Forest and hybrid models for short-term GDP forecasting in emerging economies. Asian Journal of Economic Modelling, 12(1), 33-52.

35. Winsor, C. P. (1932). The Gompertz curve as a growth curve. Proceedings of the National Academy of Sciences, 18(1), 1-8.

36. World Bank. (2025). GDP growth (annual %) - India. World Development Indicators. Retrieved from https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=IN

37. World Bank Group. (2025). India Country Overview. Retrieved from https://www.worldbank.org/ext/en/country/india

38. Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159-175.