پژوهشکده مطالعات اقتصادي و صنعتي شريف

دانشکده مدیریت و اقتصاد

پژوهشکده‌ی آمار

مرکز آمار ایران

نشست سیاست گذاران مبتنی بر شواهد

ارزیابی توانایی تکنیک ها و مدل های مختلف در پیش بینی رشد اقتصادی

تهران، ایران، ۲۰ و ۲۱ شهریور ۱۴۰۰

رضا هادی زاده

Forecasting Methods

- Linear Models:

  ARMA and ARIMA type models

  ETS models
- Machine Learning

  KNN models
- Hybrid models

  Wavelet-ARIMA

  Hybrid models

Linear Models:
Time Series Stationary Models

❖ ARMA type Models

●The ARMA (AutoRegressive Moving Average) type models are included stationary models such as AR (AutoRegressive), MA (Moving Average), ARMA and SARMA (Seasonal ARMA). The general form of ARMA (p,q) time series model is given by:

$$\phi_p(B)\,x_t = \theta_q(B)\;\varepsilon_t$$

❖ If time series process has seasonal structure represent by $SARMA(p,q)\times(P,Q)$. General form of seasonal time series is given by:

$$\phi_p(B)\,\Phi_P(B^s)\,x_t = \theta_q(B)\,\Theta_Q(B^s)\,\varepsilon_t$$

Linear Models:
Time Series NonStationary Models

❖ ARIMA type Models

●The ARIMA models are non-stationary kind of ARMA type models that converted with differencing to a stationary model. General form ARIMA models similar to ARMA model with differencing parameter that denoted by ARIMA (p,d,q) and following equation:

$$\phi_p(B)\nabla^d x_t = \theta_q(B)\ \varepsilon_t$$

❖ f ARIMA processes have seasonal structure, then modeled by

$$SARIMA(p,d,q)\times(P,D,Q)$$

General form of seasonal time series is given by:

$$\phi_p(B)\Phi_P(B^s)x_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t$$

Linear Models:
State Space Model

All ETS models are non-stationary. ETS models are considered under the umbrella first of exponential smoothing and then state-space. This type of model describes how unobserved components of the data (error, trend, and seasonality) change over time.

The general linear Gaussian state space model can be written in a variety of ways. The version used in the book is

$$y_t = Z_t \alpha_t + \epsilon_t, \quad \epsilon_t \sim N(0, H_t)$$
$$\alpha_{t+1} = T_t \alpha_t + R_t, \quad \eta_t \sim N(0, Q_t)$$

where y_t is a $p \times 1$*vector* of observations and α_t is an unobserved $m \times 1$ vector called the *state vector*. The first equations is called the *observation equation* and the second equation is called the *state equation*. The initial state vector $\alpha_1 \sim N(a_1, P_1)$. The matrices Z_t, T_t, R_t, H_t, Q_t are initially assumed to be known and the error terms ϵ_t and η_t are assumed to be serially independent and independent of each other at all time points. Matrices Z_t and T_{t-1} are permitted to depend on $y_1, y_2, \ldots y_{t-1}$. The first equation has the structure of a linear regression model where the coefficient vector α_t varies over time. The second equation represents a first order vector autoregressive process, the Markovian nature of which accounts for many of the elegant properties of the state space model.

Time series forecasting methods: ARIMA models vs ETS models

- Two of the most commonly used time series forecasting methods are ARIMA (Auto Regressive Integrated Moving Average) and ETS (Error Trend and Seasonality, or exponential smoothing).
- There are a number of factors to consider when it comes to selecting which time series forecasting method to apply, including the following:

1. what is the context of the forecast we want to do
2. what is the availability of historical data
3. the amount of accuracy/inaccuracy we can tolerate with our model
4. what time period are we forecasting for
5. what is the cost of this forecast to the company and how much value/benefit does it bring
6. how much time do we have to complete this analysis

Time series forecasting methods: ARIMA models vs ETS models

- While both methods share many similarities, below are some of the key difference between them:
- ARIMA models
  1. some are stationary
  2. do not have exponential smoothing counterparts
  3. use if you see autocorrelation in the data, i.e. the past data explains the present data well
- ETS models
  1. are not stationary
  2. use exponential smoothing
  3. use if there is a trend and/or seasonality in the data, as this model explicitly models these components

Machine Learning:
K-Nearest Neighbor Model

- KNN regression process consists of instance, features, and targets components.
- The lags parameter indicates the lagged values of the time series data. The lagged values are used as features or explanatory variables.
- This instance is used as a reference vector to find features that are the closest vectors to that instance. The relevant distance metric is calculated by the Euclidean formula as shown below:

$$\sqrt{\sum_{x=1}^{n}(f_x^i - q_x)^2}$$

- q(x) denotes the instance and f^i(x) indicates the features that are ranked in order by the distance metric. The k parameter determines the number of k closest features vectors which are called k nearest neighbors.

Hybrid Models:
WaveletArima Model

- To fit a Wavelet/ARIMA hybrid model, we follow the following steps:

1- Time series decomposition with discrete wavelet

2- Obtaining wavelet coefficients (approximation coefficients and detail coefficients)

3- Creating a approximations series and details serieas

4- Modeling the obtained series with Arima model

5- Forecasting the approximations and details Serieas

6- Making the Forecasted series from the sum of the approximation series and the details series

Hybrid Models:
Hybrid Model

Forecasting with Hybridmodels Package in R:

1- ARIMA Type model

2- ETS model

3-thetam model

The Theta model of Assimakopoulos & Nikolopoulos (2000) is a simple method for forecasting the involves fitting two θ-lines, forecasting the lines using a Simple Exponential Smoother, and then combining the forecasts from the two lines to produce the final forecast.

4- nnetar:

A feed-forward neural network is fitted with lagged values of the response as inputs and a single hidden layer with size nodes. The inputs are for lags 1 to p, and lags m to mP where m is the seasonal period specified.

Hybrid Models:
Hybrid Model

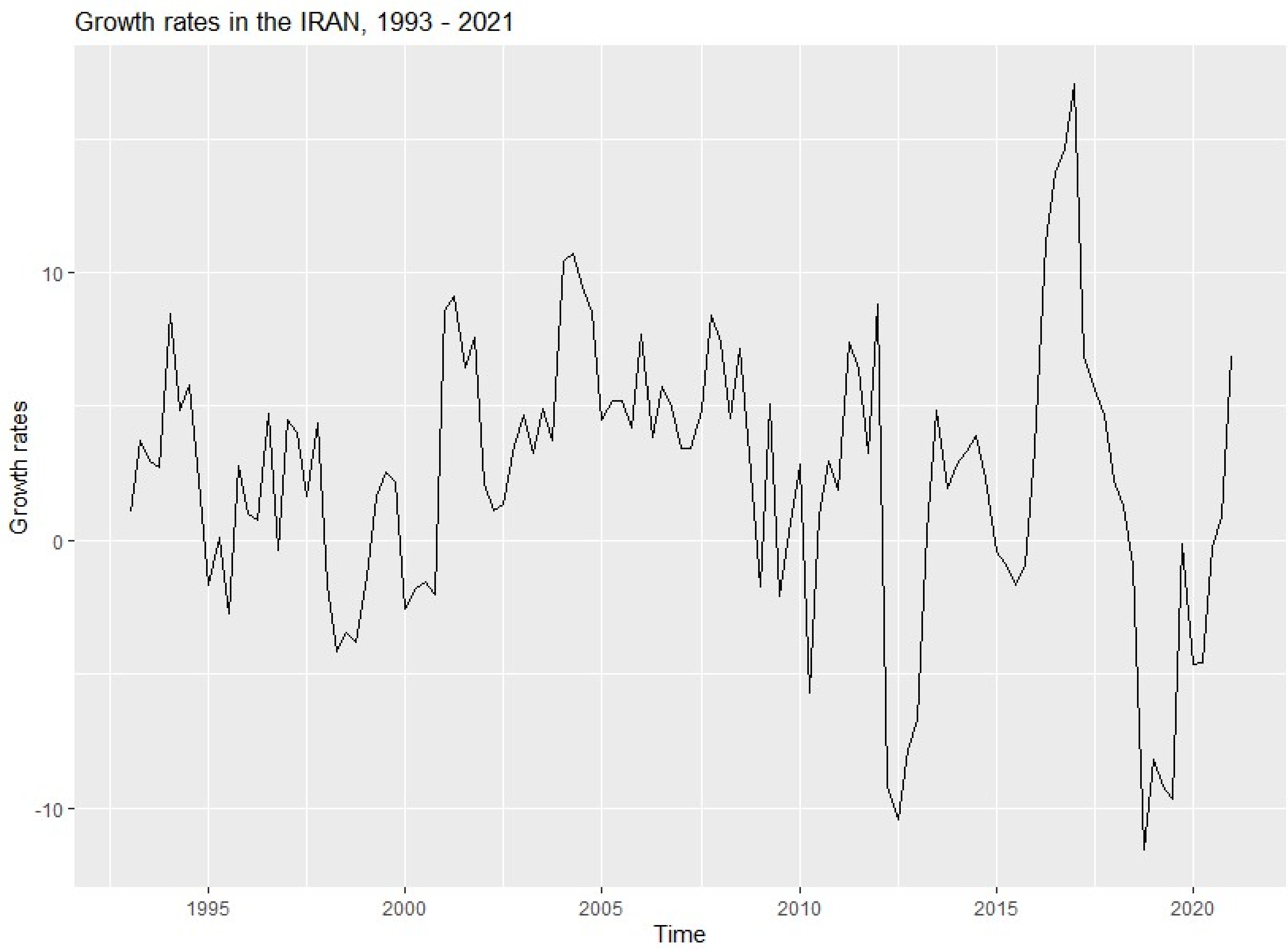
5-stlm:

takes a time series y , applies an STL decomposition, and models the seasonally adjusted data using the model passed as modelfunction or specified using method . It returns an object that includes the original STL decomposition and a time series model fitted to the seasonally adjusted data.

6- tbats model

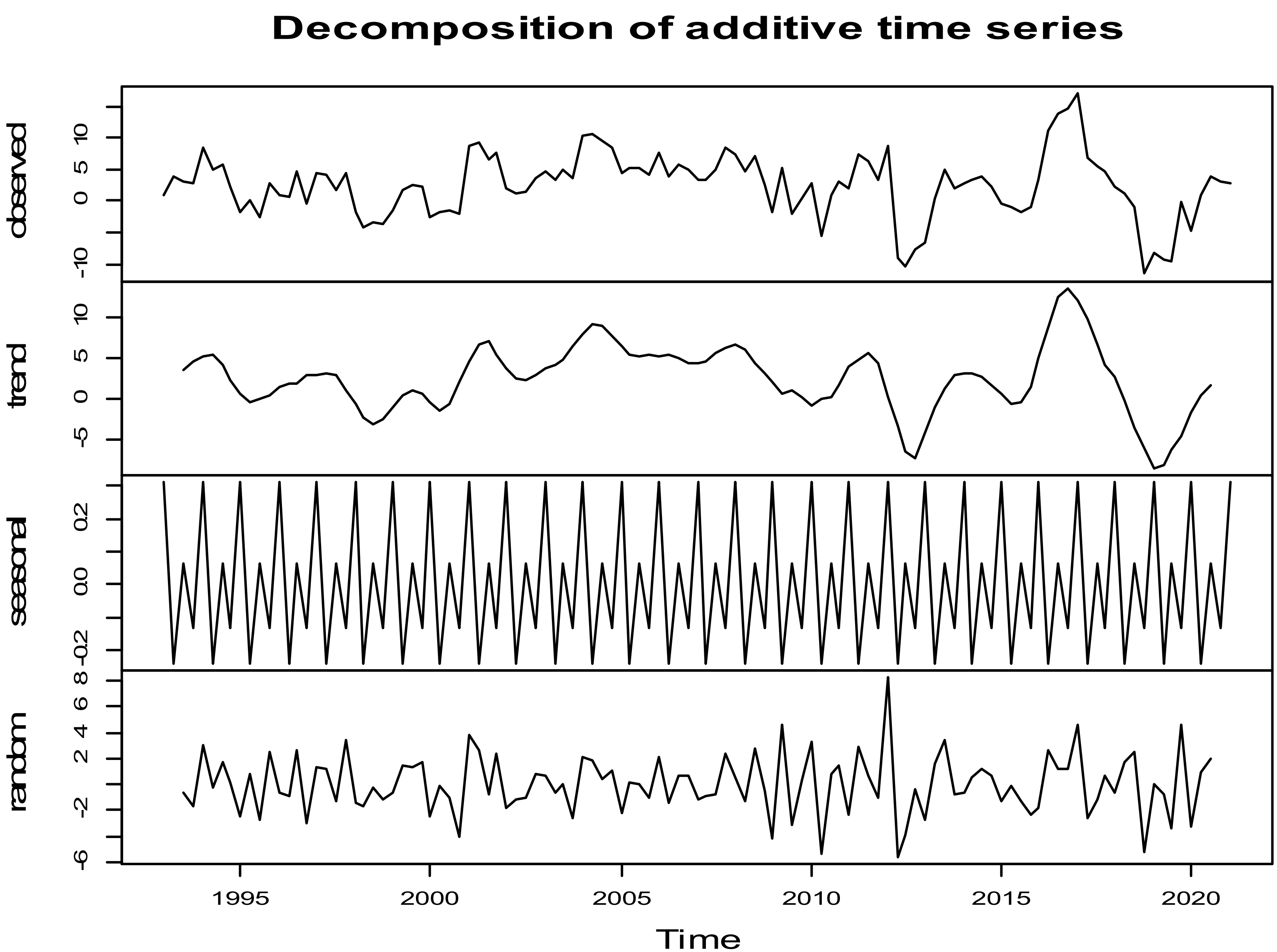
An alternative approach developed by De Livera, Hyndman, & Snyder (2011) uses a combination of Fourier terms with an exponential smoothing state space model and a Box-Cox transformation, in a completely automated manner.

Findings:



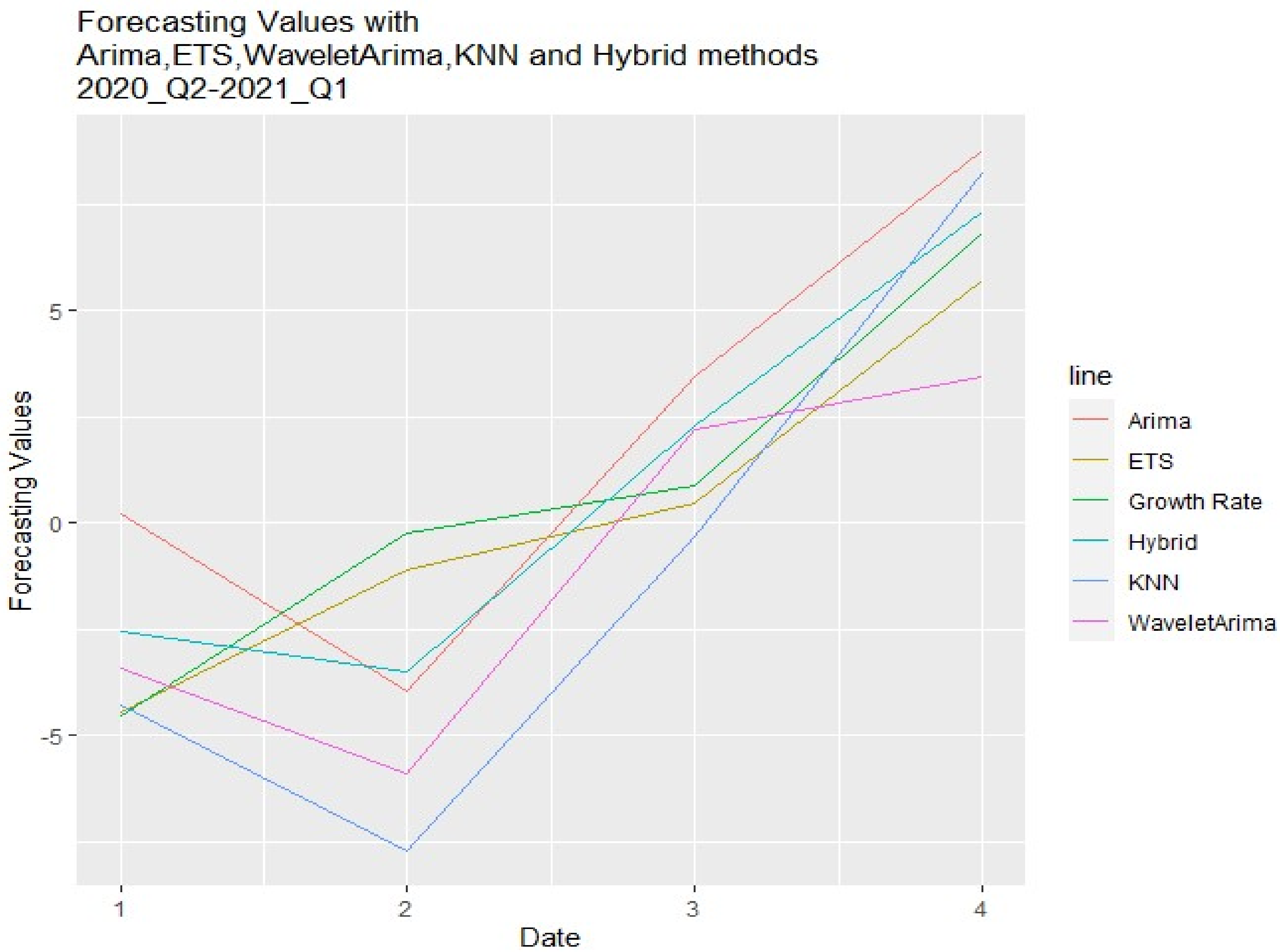
Growth rates in the IRAN, 1993 - 2021
Growth rates
10
0
-10
1995
2000
2005
2010
2015
2020
Time

Findings:



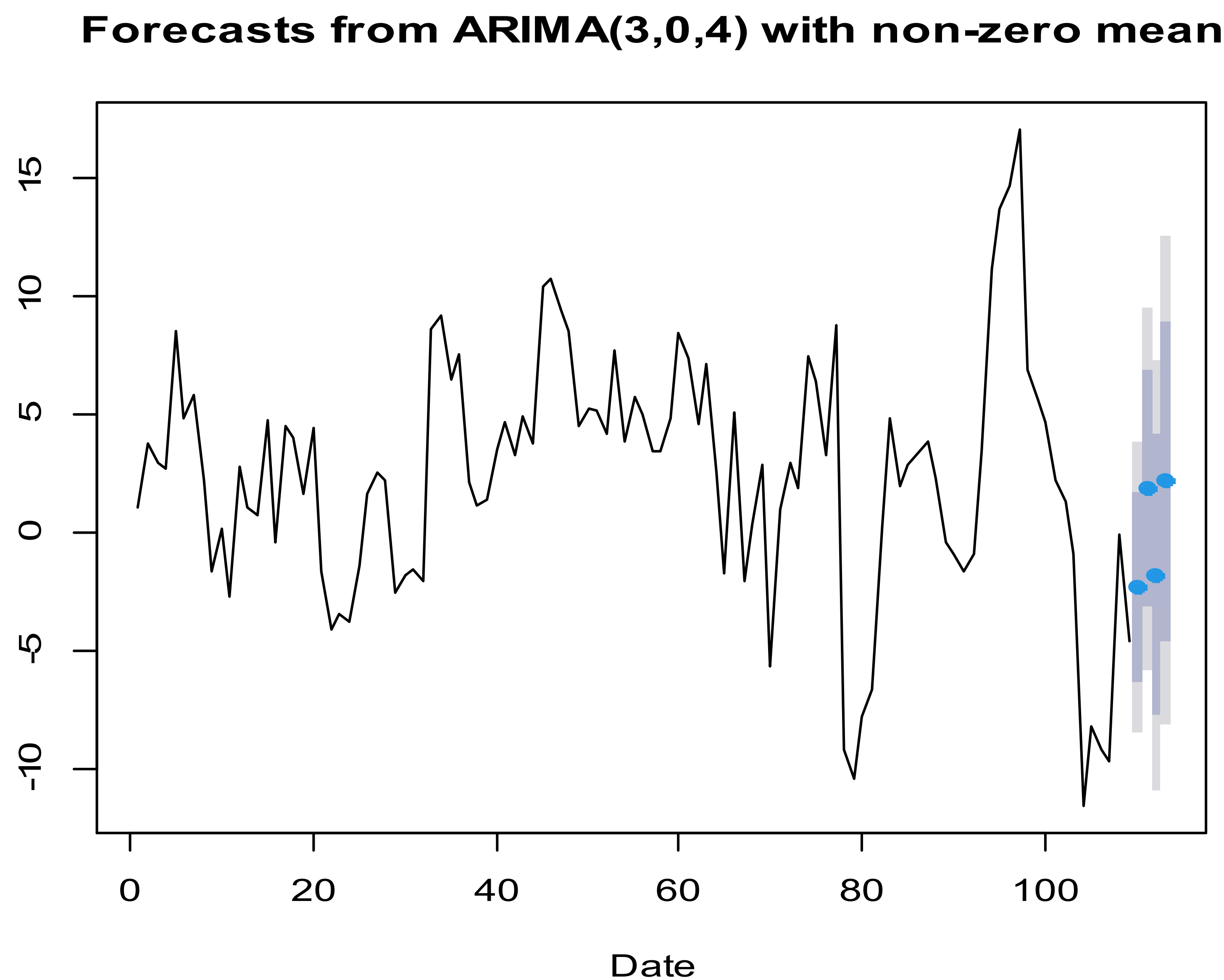
Decomposition of additive time series
observed
trend
seasonal
random
-10
0
5
10
-5
0.0
0.2
-0.2
-6
-2
2
4
6
8
1995
2000
2005
2010
2015
2020
Time

Findings:

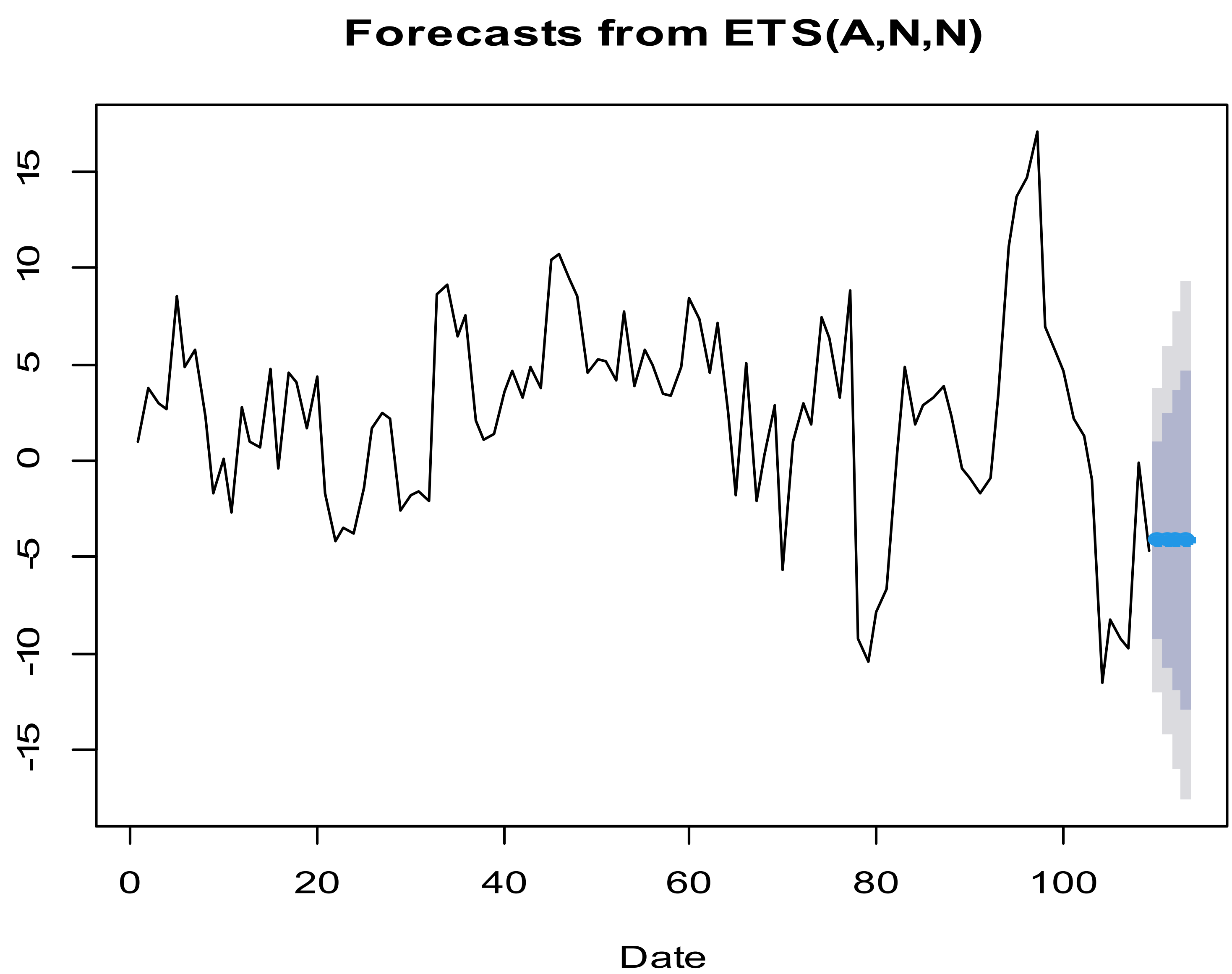


Forecasting Values with
Arima,ETS,WaveletArima,KNN and Hybrid methods
2020_Q2-2021_Q1
Forecasting Values
5
0
-5
1
2
3
4
Date
line
Arima
ETS
Growth Rate
Hybrid
KNN
WaveletArima

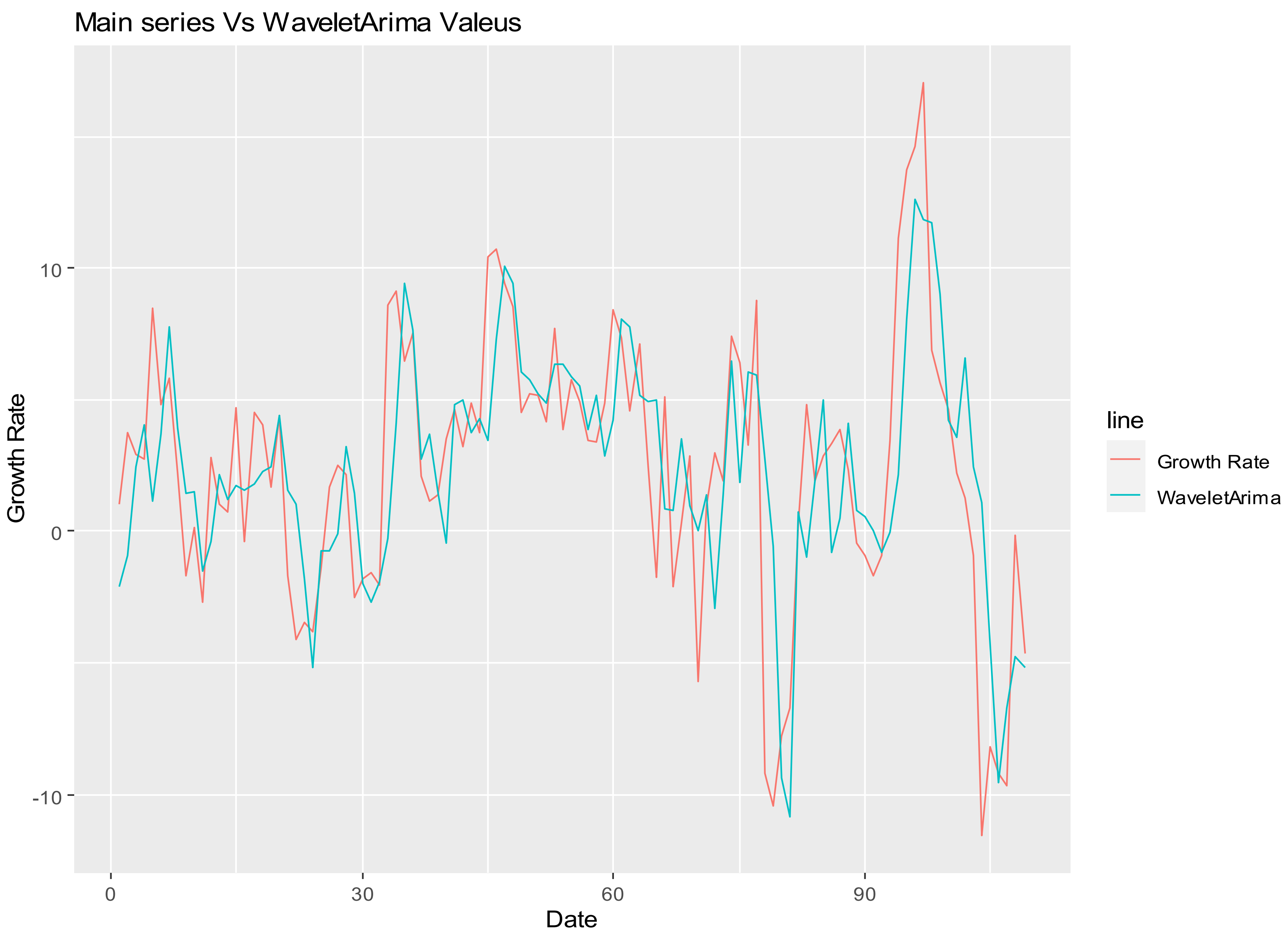
Findings:


Forecasts from ARIMA(3,0,4) with non-zero mean
15
10
5
0
-5
-10
0
20
40
60
80
100
Date

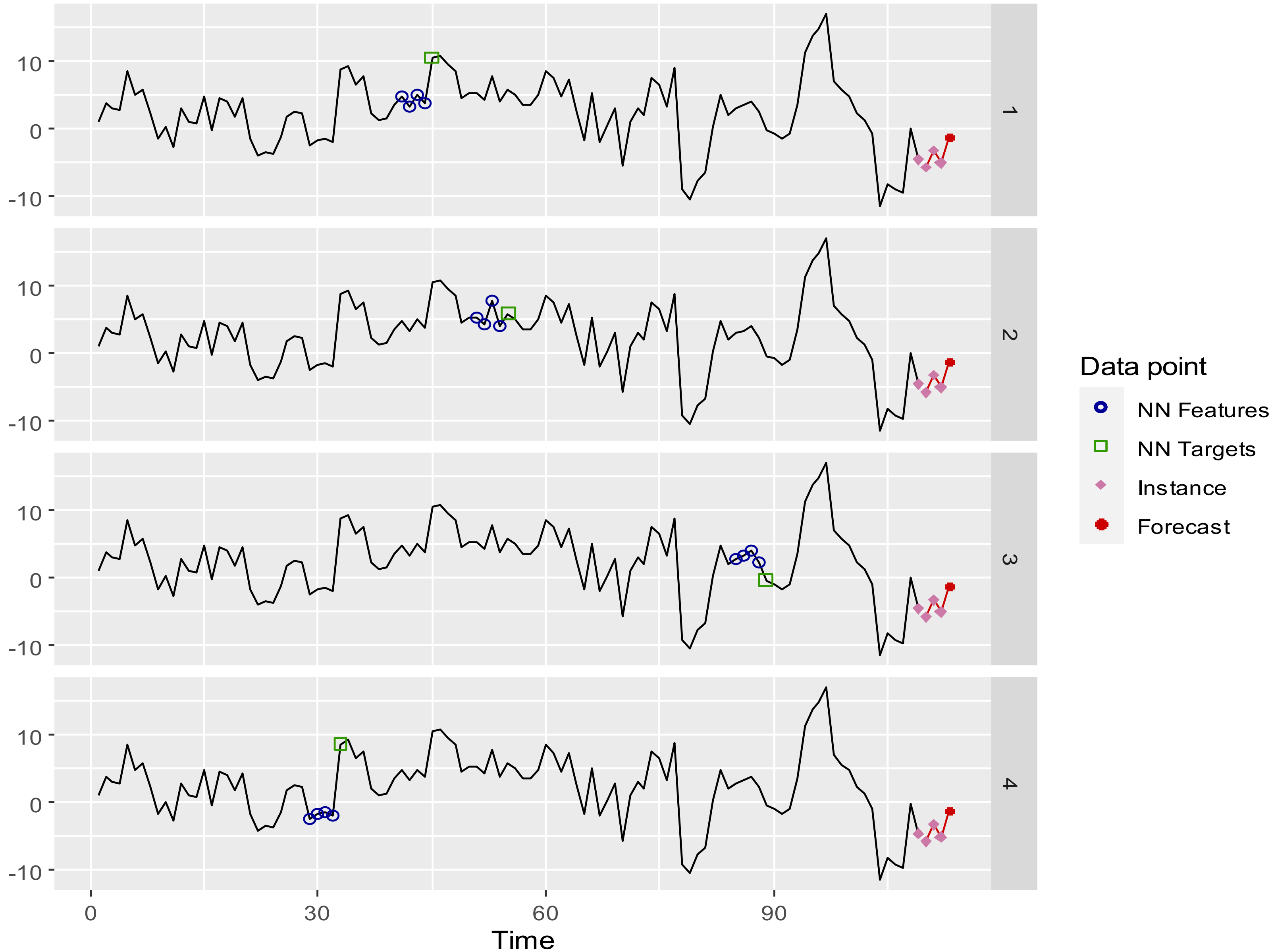
Findings:



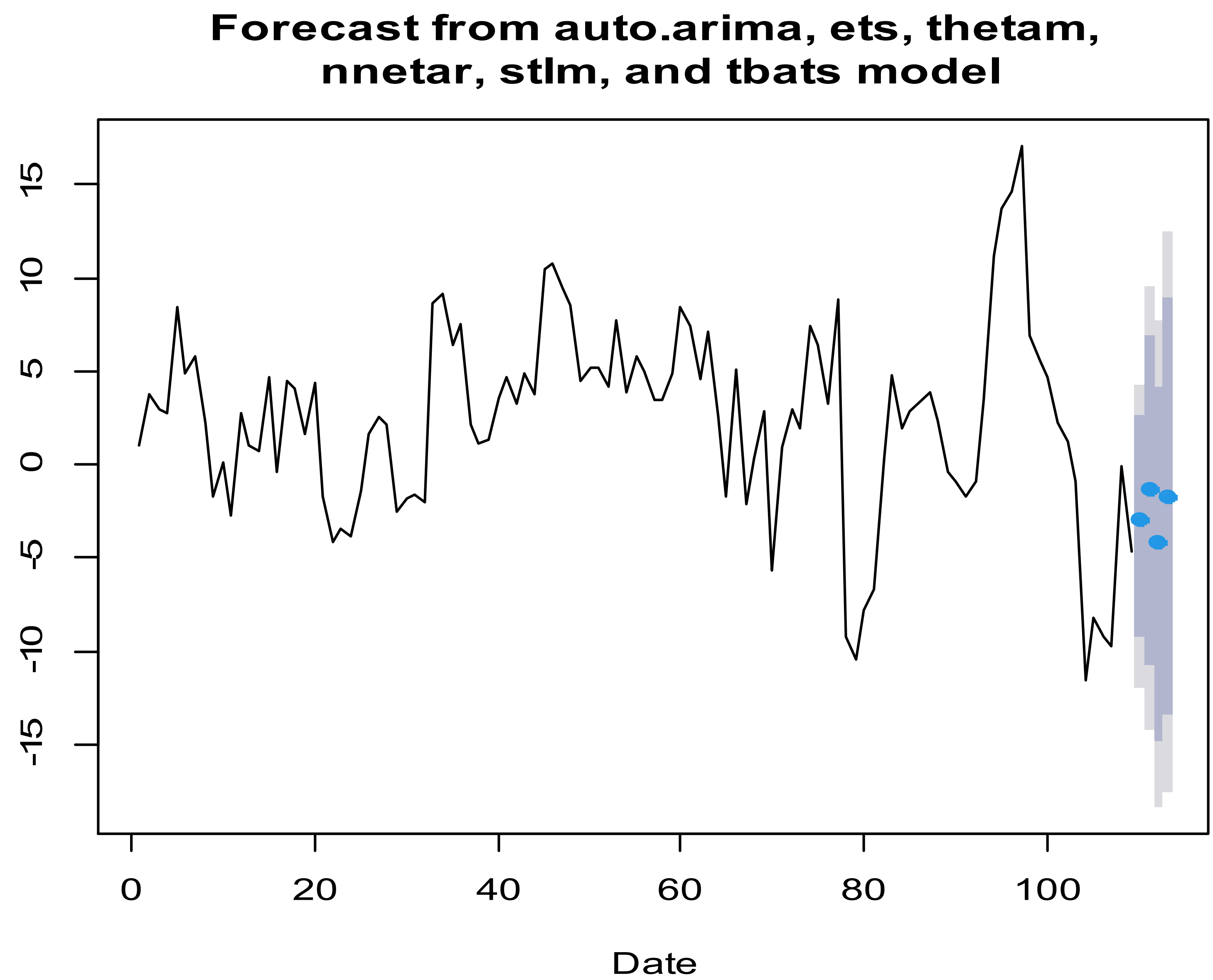
Findings:


Main series Vs WaveletArima Valeus
Growth Rate
10
0
-10
0
30
60
90
Date
line
Growth Rate
WaveletArima

Findings:


10
0
-10
1
2
3
4
0
30
60
90
Time
Data point
NN Features
NN Targets
Instance
Forecast

Findings:


Forecast from auto.arima, ets, thetam, nnetar, stlm, and tbats model
15
10
5
0
-5
-10
-15
0
20
40
60
80
100
Date

Thank you!