Technical Disclosure Commons

Defensive Publications Series

November 2021

TIME SERIES PREDICTIVE MODEL FOR PROVISIONING OF CLOUD RESOURCES

HP INC

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation
INC, HP, "TIME SERIES PREDICTIVE MODEL FOR PROVISIONING OF CLOUD RESOURCES", Technical Disclosure Commons, (November 28, 2021)
https://www.tdcommons.org/dpubs_series/4742

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

TIME SERIES PREDICTIVE MODEL FOR PROVISIONING OF CLOUD RESOURCES

ABSTRACT

Predictive modeling or Forecasting is to predict the future using (time series related or other) data we have in hand. Cloud Infrastructure as a solution facilitates pay-as-you-use execution environments that scale transparently to the user. There is continuous challenge of configuring cloud infrastructure to provide maximal performance while minimizing the cost of resources used. This disclosure focuses on time series data modeling to predict business requirements of a Cloud Platform and its application to daily infrastructure management.

PROBLEM STATEMENT

Actual objective of cloud infrastructure is to flex up and flex down based on business demand to save cost. However, if the cloud infrastructure capacity for the Cloud services is fixed, load must be checked based on yearly aggregate and extend the capacity during year end if needed. This does not substantiate the **cost savings** on the infrastructure. Often the capacity is **overly estimated** due to fear of failing during peak loads. This becomes an impediment if the infrastructure is fixed for each service and if the load is not shared across different services based on the usage. This disclosure provides a **predictive model** for scaling, scheduling, and planning of Cloud capacity requirements.

BRIEF DESCRIPTION OF THE DRAWINGS

The drawings are provided to illustrate the disclosure as explained in the detailed description

Fig 1 provides Cloud resources usage trends of a year

Fig 2 provides Data Forecasts from ARIMA (5,1,1)

Fig 3 provides Actual data trends and Forecasts from ARIMA (5,1,1)

SOLUTION

A predictive statistical data model for business forecasting to plan daily capacity on Cloud Infrastructure for a time frame (daily/monthly/quarterly) and auto flex the servers. Statistical model is applied on **R platform** with different statistical methods for data modeling. The solution would predict/forecast the cloud resources usage, cloud computing utilization data, server processing usage.

Data Decomposition

1. Estimate the trend with two different approaches with a **smoothing procedure** such as moving averages without using equations or model the trend with regression equation.
2. Seasonal factors are estimated with **monthly data** which entails estimating an effect for each month of the year. For **quarterly data**, this entails estimating an effect for each quarter.
3. Data is modelled with an **ARIMA model**. When deep looking into seasonal component, there is high correlation on day of the week trend. Seasonal patterns repeat consistently throughout time frame. Thereby, a trend is estimated using time series data modelling. For the decreasing trend having fixed infrastructure and not scaling down would incur huge cost. Predictive Modeling is done for load of Cloud resources. Cloud resources data is being tracked for one year (January to December data) as shown in Fig 1. Hence data time frame chosen or available for modeling was one year. This is considered as training dataset. 2 months data is chosen as hold-out sample. Hold out sample is called test dataset. This is to predict and compare the actual vs predicted values for 2 months and measure the error as **Mean absolute percentage** error (MAPE). We try to fit a model and capture the residuals or error out of the model. Once the residuals pass white noise test, preferred model is chosen. With application of different models and finally arrived at a model with least Mean Absolute Percentage error.

Different Statistical models can be chosen based on data trends with better forecasted data as a best fit model.

Application of Models with Results:

Exponential Smoothing Model

Smoothing techniques can be applied. Simple Exponential method on data gave metrics with **Mean absolute percentage error** (MAPE) as **13.44.** Box pierce test was conducted on the residuals of the fit to see if it passes the white noise test. White noise test passes if p value is greater than 0.05. When it passes white noise test, it will indicate that the fit of the model is good, and residuals is just some noise which can't be further predicted. Box test p value is greater than 0. 05.

- The forecast at time t+1 is equal to a weighted average between the most recent observation yt and the most recent forecast y^t|t−1

$$\hat{y}_{t+1|t} = \alpha y_t + (1-\alpha)\hat{y}_{t|t-1}$$

ARIMA Model: Autoregressive Integrated Moving Average Model

- **ARIMA(p,d,q)** is the description for the model where parameters *p*, *d*, and *q* are -negative integers, *p* is the order (number of time lags) of the autoregressive model, *d* is the degree of differencing (the number of times the data have had past values subtracted), and *q* is the order of the moving-average model.
- Seasonal ARIMA models are usually denoted **ARIMA(*p*,*d*,*q*)(*P*,*D*,*Q*)*m***, where *m* refers to the number of periods in each season, and the uppercase *P*,*D*,*Q* refer to the autoregressive, differencing, and moving average terms for the seasonal part of the ARIMA mode

Different values of p and q are tried and arrived at **ARIMA(5,1,1**) as best model with least error**.**

IMPLEMENTATION OF SOLUTION

This framework is applied on Cloud Infrastructure for the workload of a year. Resource usage pattern of a year is shown in Fig 1. ARIMA is chosen as best fit model. Predicted data for ARIMA(5,1,1) is shown in Fig 2. Prediction intervals done with 95% confidence interval for forecasting. The Predicted data in blue color is similar to the actual data in Fig 1. Actual data and Predicted/Forcasted data from ARIMA(5,1,1) is shown in Fig 3.

PRIOR SOLUTIONS

There are cloud providers having ASG (Auto Scaling Group) which does auto-scaling based on minimum/maximum configuration of nodes. Auto Scaling solution is relevant while consuming that cloud-provider only.

USE CASES

- Use Case 1: Application of time series prediction models can be used for prediction of infrastructure usage of cloud services.
- Use Case 2: Forecasting lies within the prediction's 95% confidence interval. Thereby, we can decide to adjust the forecast for the fiscal year or take other actions that mitigate the forecasting gap.
- Use Case 3: Use Predictive planning to predict future performance based on your historical data. Prediction helps to understand the past behaviour/patterns for future predictions for resource usage.
- Use Case 4: Used to compare and validate plans and forecasts based on the predictions. Such analysis helps to identify the factor that influences the fluctuation in data usage.
- Use Case 5: Helps to compare the historical predictions against the historical forecasts to determine the accuracy of each.
- Use Case 6: With historical and actual data, we can gauge how far the forecast and the predicted data had diverged from the actual data.
- Use Case 7: Performance analysis of time series on resources helps to compare the present performance of the services with that of the past.
- Use Case 8: Helps in cost estimation of resources needed, thereby optimizing business expenses.
- Use Case 9: Historical data patterns can be used to predict future product behavior.

ADVANTAGES

- **Resource predictions** - Time series prediction helps to understand the past behavior/patterns for future predictions for cloud resource usage or cloud computing utilization data. Trend is estimated using time series data modelling
- **Data usage** - Such analysis helps to identify the factor that influences the fluctuation in data usage.
- **Performance analysis** - The analysis of time series on resources helps to compare the present performance of the services with that of the past.
- **Data analysis** - Time series techniques on resources helps to analyze data.
- **Budget estimation** – Optimization of business expenses.

- **Product sales analysis** - Contributes to sales and product analysis.
- **Product behavior** - Historical data patterns can be used to predict future product behavior.
- **Reliable solution** - Reliable solution as it is time series dependent.
- **Time variant analysis** - Information can be extracted by measuring data at various time intervals such as hourly, daily, monthly, quarterly, annually or at any other time interval.

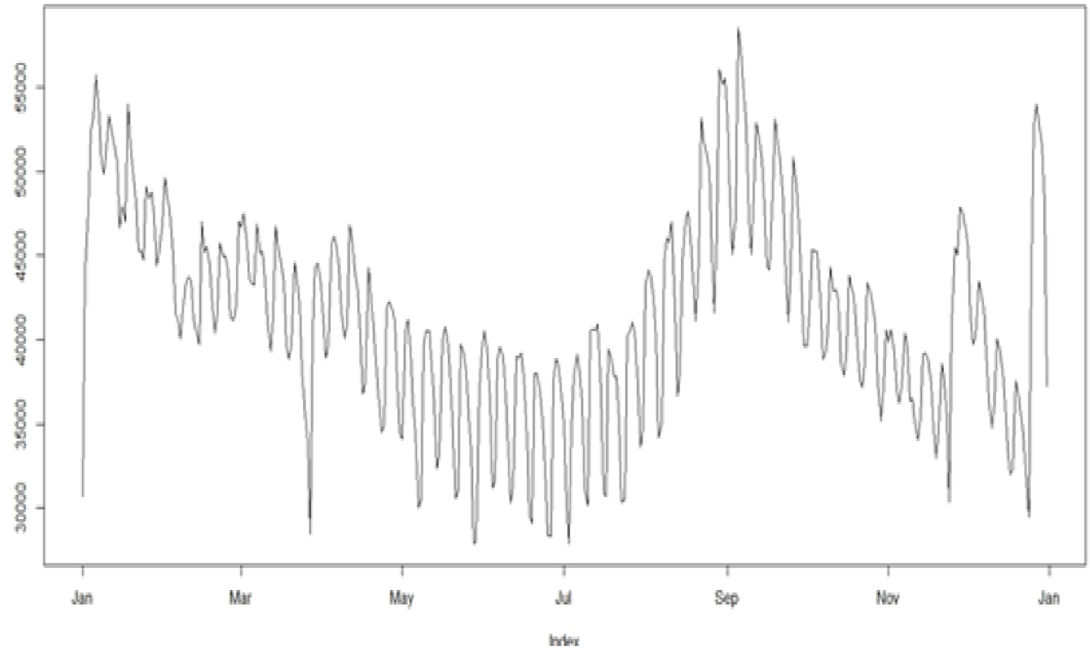

30000
35000
40000
45000
50000
55000
Jan
Mar
May
Jul
Sep
Nov
Jan
Index

Fig 1

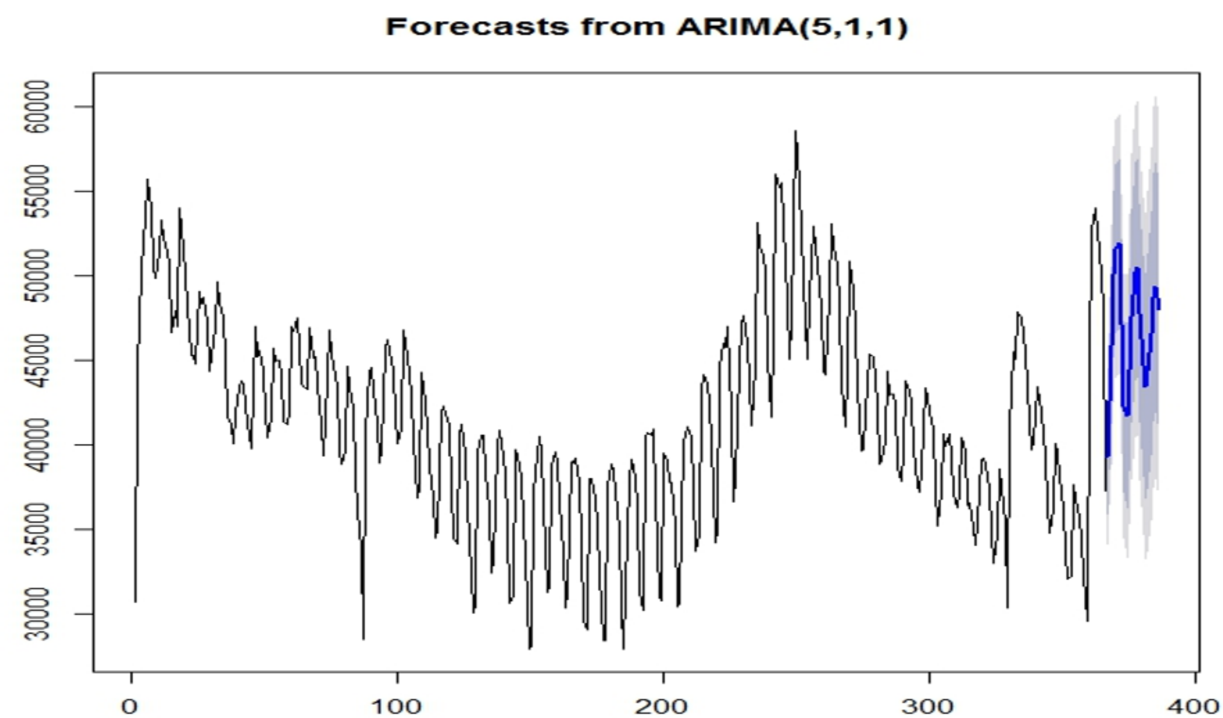
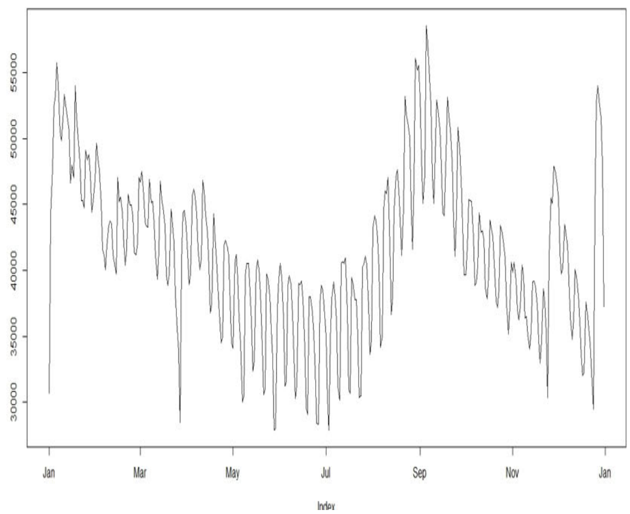

Forecasts from ARIMA(5,1,1)
30000
35000
40000
45000
50000
55000
60000
0
100
200
300
400

Fig 1



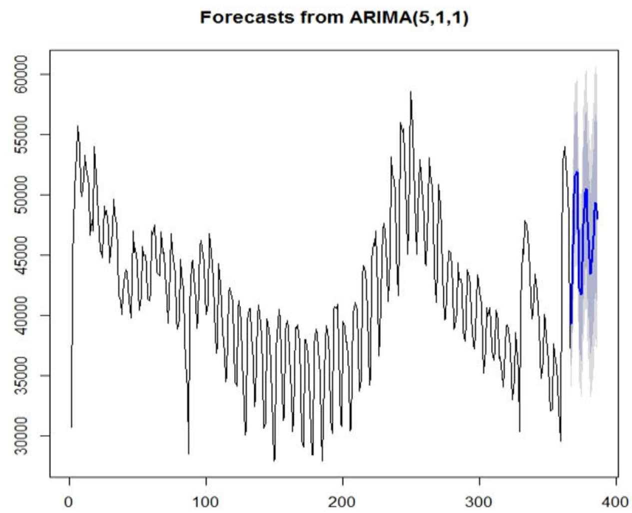


Fig 3

Disclosed by Pushpalatha K R, HP Inc.