The Impact of Lengths of Time Series on the Accuracy of the ARIMA Forecasting

Lu Qin[1], Kyle Shanks[1], Glenn Allen Phillips[1] & Daphne Bernard[1]

[1] Institutional Research and Assessment, Howard University, Washington, D.C., USA

Correspondence: Lu Qin, Institutional Research and Assessment, Howard University, Washington, D. C., USA. Tel: 1-202-865-0206.

Received: September 10, 2019 Accepted: September 27, 2019 Online Published: September 29, 2019

doi:10.5430/irhe.v4n3p58 URL: https://doi.org/10.5430/irhe.v4n3p58

Abstract

The Autoregressive Integrated Moving Average model (ARIMA) is a popular time-series model used to predict future trends in economics, energy markets, and stock markets. It has not been widely applied to enrollment forecasting in higher education. The accuracy of the ARIMA model heavily relies on the length of time series. Researchers and practitioners often utilize the most recent 1- to 5-years of historical data to predict future enrollment; however, the accuracy of enrollment projection under different lengths of time series has never been investigated and compared. A simulation and an empirical study were conducted to thoroughly investigate the accuracy of ARIMA forecasting under four different lengths of time series. When the ARIMA model completely captured the historical changing trajectories, it provided the most accurate predictions of student enrollment with 20-years of historical data and had the lowest forecasting accuracy with the shortest time series. The results of this paper contribute as a reference to studies in the enrollment projection and time-series forecasting. It provides a practical impact on enrollment strategies, budges plans, and financial aid policies at colleges and institutions across countries.

Keywords: ARIMA, time series, forecasting, simulation, empirical analysis, enrollment prediction

1. Introduction

Time-series analysis has been used successfully to predict future changes in economics, energy markets, and stock markets (Wang, Hsu, & Lion, 2011; Zhang, 2003; Ward, 2007; Chen, 2008). Accurately predicting how many jobs are needed, how stocks will increase or decrease, and how oil prices will fluctuate has a significant impact on trade and business. Consequently, the accuracy of the time-series analysis is a significant focus of forecasting research due to its impact on investment and budget planning.

In enrollment forecasting, multiple factors influence projection accuracies such as the forecasting method, length of time series, frequency of enrollment occurrence, and other institution-related predictors (e.g., enrollment policy, financial aid policy, and economic situation). Two traditional approaches to enrollment forecasting are causal modeling and time-series analysis (Brinkman & McIntyre, 1997). The causal modeling approach predicts future enrollment, regressing on factors that affect enrollment levels such as financial aid policy, students' demographics (e.g., gender, ethnicity, age), employment status, and prior educational experience (Brinkman & McIntyre, 1997). This approach is challenging to apply to the context of higher education as these types of data are difficult to collect and are sometimes highly related to each other. The multicollinearity among predictors in causal modeling leads to a biased estimate of the enrollment outcome (Peng, Lee, & Ingersoll, 2002). Time-series analysis, as opposed to traditional approaches, predicts future enrollments by tracing the trajectories of historical enrollment data. It requires fewer predictive factors and is easier to understand and implement.

The Autoregressive Integrated Moving Average model (ARIMA) is a popular time-series model used to predict future trends with high accuracy (Zhang, 2003; Contreras, Espinola, & Nogales, 2003; Conejo, Plazas, & Espinola, 2005). This study focuses on exploring the impact of lengths of time series on the accuracy of prediction using the ARIMA model.

The rest of the paper is organized as follows. Section 2 reviews the ARIMA model forecasting based on the time-series data. Next, a simulation study is illustrated in section 3. In section 4, an empirical study is reported. Finally, section 5 includes the results of the two studies and discusses the implications for future research.

2. Literature Review

2.1 ARIMA Model

The ARIMA model is an integrated model that combines an autoregressive (AR) model and a moving average (MA) model in time-series prediction based on the Box-Jenkins methodology (Box, Jenkins, & Reinsel, 2015). The prediction in the ARIMA model for a stationary or non-stationary time series is a linear function that includes lags of dependent variables (future values at specific time points) and lags of random errors. In a stationary time series, means and variances remain constant across time points without linear or nonlinear trends (Ward, 2007). The difference between the ARIMA model and the AR or MA model is that the ARIMA model is designed to analyze a non-stationary time series by differing the autocorrelations among lags of time series (Wang, Hsu, & Liou, 2011). Let $\boldsymbol{Y} = (y_1, y_2, \dots, y_t)$ be a vector of values at time series $\boldsymbol{TS} = (1, 2, \dots, t)$, μ be the conditional means, $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t)$ be a vector of random errors, and $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_t)$ be the coefficients in the linear function. If we let p be the number of autoregressive terms in the AR model, the AR model can be expressed as:

$$y_t = \mu + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p} + \varepsilon_t \tag{1}$$

Here, y_t is the predicted value at time point t and ε_t is the random error at time point t, which falls in a normal distribution with a mean of 0 and a variance of σ^2. $\boldsymbol{\beta}$ represents the strength of the relationship between the value at a previous time point and the predicted value at a future time point. The same methodology applies to the moving average (MA) model. If we let q be the number of lagged errors in the MA model, the predicted value y_t is regressed on past errors as:

$$y_t = c + \vartheta_1 \varepsilon_{t-1} + \vartheta_2 \varepsilon_{t-2} + \cdots + \vartheta_q \varepsilon_{t-q} + \varepsilon_t \tag{2}$$

Similarly, ϑ represents the strength of relationships between a measurement error at a previous time point and the predicted value at a future time point. Therefore, an autoregressive integrated moving average (ARIMA) model would be written in one equation as:

$$y_t = \mu + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p} - \vartheta_1 \varepsilon_{t-1} - \vartheta_2 \varepsilon_{t-2} - \cdots - \vartheta_{t-q} \varepsilon_{t-q} \tag{3}$$

Three parameters are included in the ARIMA model: p, q, and d. Parameters p and q refer to the numbers of autocorrelation and lagged random errors, respectively, while d represents the level of changes across times (Chen, 2008). When past data follows a non-stationary changing trajectory, such as an upward trend, a downward trend, or a seasonal trend, the ARIMA model includes the third parameter d to account for differences and to adjust for stationarity in time series. When $d = 1$, the difference is expressed as $y_t - y_{t-1}$, and when $d = 2$, the difference is expressed as $(y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$, and so forth.

Based on three parameters, the ARMIA (p, d, q) model captures at least six types of changing trajectories. ARIMA (1,0,0) is the first-order autoregressive (AR) model that accounts for the first autocorrelation among past time series. ARIMA (0,1,0) is the random walk model, which adjusts differences for stationarity. ARIMA (0,0,1) is the one-lagged moving average (MA) model, in which the future value is regressed on the one lagged random error. ARIMA (1,1,0) is the differenced first-order autoregressive model that consists of one autocorrelation as well as adjustments of seasonal differences. ARIMA (0,1,1) is the differenced one-lagged moving average model that consists of one lag of random error and differences adjustments for non-stationarity. ARIMA (1,1,1) is the integrated model that includes a first-order autoregressive model, one lagged moving average model, and non-stationarity adjustments. The number of p, q, and d parameters is arbitrary and depends on the length of past time series, historical data trends, and autocorrelation among time series. When a long sequence of time series is involved, such as 20- or 30-years of data, the parameters p, q, and d need to be larger to sufficiently cover the autocorrelation across times. By adjusting the p, q, and d parameters, the ARIMA model captures previous data trajectories with high accuracy.

2.2 Previous Studies

The ARIMA model has been widely applied in fields such as economics, energy, and finance. Further, the accuracy of the short-term prediction of the ARIMA model has been validated in a variety of contexts (Contreras, et al., 2003; Morana, 2001; Gross & Galiana, 1987). Contreras et al. (2003) accurately predicted next-day electricity usage with the ARIMA model. In another study, Morana (2001) applied the ARIMA model to predict the short-term oil price. Gross & Galiana (1987) also successfully used the ARIMA model to forecast short-term loads. In fact, the superior performance of the ARIMA model over other linear time-series models, such as AR, MA, ARMA, has been documented in many energy forecasting studies (Valipour, Banihabib, & Behbahani, 2013; Marriott & Newbold, 1998; Chujai & Kerdprasop, 2013; Colak, Yesilbudak, Genc, & Bayindir, 2015; Wang, Hsu, & Liou, 2011). Additionally, the ARIMA model was found to have fewer residual errors than the ARMA model (Valipour et al., 2013) and less bias than either MA or AR model (Wang, Hsu, & Liou, 2011).

The application of the ARIMA model in higher education, however, has not been fully investigated. O'Bryant (1991) first introduced the ARIMA model in enrollment forecasting at Sinclair Community College and found that the ARIMA model accurately predicted budget expenses for the University. Brinkman and McIntyre (1997) compared traditional forecasting models (e.g., regression model) and the ARIMA model. They found that the time-series model did not need as much data as the traditional regression model.

2.3 Statement of the Problem

The accuracy of the ARIMA model heavily relies on the length of time series (Box, Jenkins, & Reinsel, 2015). In fact, projection accuracy increases concurrently with the length of the time series (Valipour, Banihabib, & Behbahani, 2013; Zhang, 2003; Wang et al., 2011). Chen (2008) suggested using 45 to 60 years of past data to achieve a good prediction accuracy in the ARIMA model. Most institutions, however, only keep up to 30 years of historical student data in their student reporting database. Therefore, including 45- to 60-years of historical data in the ARIMA model is not realistic for enrollment projection in higher education. Institutional researchers and practitioners most often utilize the most recent 1- to 5-years of historical student data to predict future enrollment. The accuracy of enrollment projection with a 5-year series of historical data, a 10-year series of historical data, a 20-year series of historical data, and a 30-year series of historical data has never been investigated and compared.

No previous study has focused on the impact of length of time series on the projection accuracy in enrollment forecasting using the ARIMA model. This study aims to provide a thorough investigation of differences in the projection accuracy under different conditions of the length of time series using the ARIMA model.

3. A Simulation Study

A simulation study was conducted to evaluate the impact of the length of time series on the forecasting accuracy of the ARIMA model.

3.1 Data Generation

Undergraduate student enrollment data from the top 10 Historically Black Colleges/Universities (HBCUs) were downloaded from the Integrated Postsecondary Education System (IPEDS) national database. The RStudio environment (Studio, 2012) was used to simulate enrollment data for four conditions corresponding to the length of the time series: 5-, 10-, 20- and 30-years. The enrollment data for the four conditions were simulated from a normal distribution with a mean of μ and a variance of σ, where μ and σ were the average of the top 10 HBCUs undergraduate enrollment and standard deviation, respectively.

$$y_t \sim N(\mu, \sigma) \tag{4}$$

A vector of 10 data was replicated 100 times under the first condition, in which the first 5 data were treated as the true future enrollment that was used to compare with the predicted values in the ARIMA model, and the 6^{th} to 10^{th} data were treated as the historical data and was used to forecast in the ARIMA model. Likewise, a chain of 6^{th} to the 15^{th} vector, a chain of 6^{th} to the 25^{th} vector, and a chain of 6^{th} to the 35^{th} vector was also replicated 100 times under the 2^{nd}, 3^{rd}, and 4^{th} simulation conditions. The first five data were used as the true future enrollment, and the rest of the data was used as previous time-series data for forecasting. Therefore, a $5 \, x \, 100$ data matrix, a $10 \, x \, 100$ data matrix, a $20 \, x \, 100$ data matrix, and a $30 \, x \, 100$ data matrix were used in the ARIMA model for time-series forecasting.

3.2 Analysis

The Tseries package (Trapletti & Hornik, 2018) in RStudio (Studio, 2012) was used to capture historical trajectories and forecast enrollment. The unique p, d, q parameters of each replication were identified to recover past trajectories. The parameter p (AR) and the parameter q (MA) were identified by checking the partial autocorrelation function (PACF) and the autocorrelation function (ACF). The lags where the PACF cut off indicated the number of p and the lags after where ACF cuts off indicated number of q. For example, a first-order MA model is characterized as a large autocorrelation at lag $q = 1$ and the partial autocorrelations diminish to zero after the first lag (Kinney, 1978). The d parameter (stationarity of time-series data) was checked by the Augmented Dickey-Fuller test (ADF) as following,

$$ADF = \frac{\rho}{SE(\rho)} \tag{5}$$

Where ρ is the estimated coefficients of time-series data. Statistically significant results $(p \leq 0.05)$ suggest that the time-series is stationary. If the time-series data is non-stationary, the difference (d parameter) would be added in the ARIMA model to adjust the auto-correlation between lags.

The ARIMA model uses the maximum likelihood algorithm (MLE) to estimate parameters. The maximum likelihood estimation identifies parameter values that maximize the log-likelihood of the data (Enders, 2010). When the MA model is not involved, the ARIMA model is simplified as a linear autoregression model, in which the ordinary least squares (OLE) algorithm is used to estimate parameters. OLE is used to identify parameter values that minimize the squared residual errors of the predicted values (Keith, 2014). In this study, MLE was used in the analysis to estimate parameters.

Multiple indices were used to assess model fit: recovery rate, the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), and Mean Squared Errors (MSE). The recovery rate was assessed by comparing the true trajectories and the trajectories proposed by the model. When the full recovery of the trajectory was achieved, the best model was the one with the lowest AIC, BIC, and MSE indices. The bias of the prediction, an essential criterion to determine the performance of the ARIMA model, was calculated by the average of differences between the forecasted values and the simulated "true" data across 100 replications.

$$Bias = \frac{\sum |\left(y_t^{predicted} - y_t^{observed}\right)|}{Replications} \tag{6}$$

3.3 Results

(a) ADF test

The Augmented Dickey-Fuller Test (ADF) showed that the time-series data were non-stationary $(p \geq 0.05)$ across 100 replications under four conditions ($\boldsymbol{TS} = 5, 10, 20, 30$) in the simulation. The difference parameter, d, was added in the model to adjust the autocorrelation between lags.

(b) ACF and PACF Test

The example ACF and PACF plots are shown in Figure 1. The parameter p for the AR model varied between 2 to 4 when the time series was 5, varied between 5 and 8 when the time series was 10, varied between 12 and 16 when the time series was 20, and varied between 14 and 23 when the time series was 30 across 100 replications. The parameter q for the MA model varied between 1 and 3 when the time series was 5, varied between 3 and 7 when the time series was 10, varied between 1 and 11 when the time series was 20, and varied between 0 and 15 across 100 replications when the time series was 30.

(c) Model Fit

Four different types of ARIMA model were fit in the analysis: AR model (1,0,0), MA model (0,0,1), ARMA model (1,0,1), and differenced integrated model (1,1,1). By fitting the different values of p, d, and q in the ARIMA model in each replication under four conditions of the length of time series, the selected model was the one that maximally recovered the past changing trajectories.

One replication's model fit is shown as an example in Figure 2. The solid red line indicates the changing trajectories of the true simulated data under different lengths of time series. The blue dashed line is the pattern captured by the ARIMA model. The blue line recovered the true data changing trajectories under four lengths of time series across 100 replications. The high model-fit rates indicated that the ARIMA model was able to provide accurate forecasting based on the past changing patterns.

(d) Forecasting

The best-fitting model was then used to forecast future time-series data. As shown in Figure 3, the black lines indicate the past data trajectories, and the blue lines indicate the 10 forecasting values and its changing trajectories. The impact of length of time series on the forecasting accuracy was evaluated by comparing the forecasting values and the simulated "true" values across 100 replications under four conditions of the length of time series as shown in Table 1.

One of the findings that were consistent with the earlier studies was that the most considerable bias between the predicted values and the simulated "true" data was found in the shortest time series ($TS = 5$). The sequence of time series at $10, 20,$ and 30 provided significantly higher accuracy in forecasting than the sequence of time series at 5 ($F = 6.8, p < .001$). Contrary to prior research, where the most extended time series provides the most accurate forecasting, results in this study indicated that when the ARIMA model fit the past data the sequence of the time series at 20 provided the smallest bias between the forecasted values and the simulated "true" data ($bias = 19.9$); however, there were non-significant differences among the sequences of time series at $10, 20,$ and 30 ($p > .05$).

This result suggests two important implications. First, when the ARIMA model completely fit the past data, the longest time series might not be the one to provide the most accurate forecasting. Second, if the sequence of the time series at 20 is not achievable under any circumstances, the sequence of the time series at 10 would be able to provide a prediction with an acceptable accuracy; however, the sequence of the time series at 5 would be too short to provide a relatively accurate prediction.

4. An Empirical Study

We applied information gleaned from the simulation study to a real data set to explore the impact of the length of time series on the accuracy of an enrollment projection using the ARIMA model. A total of 35 years of fall term enrollment data of undergraduate students at Howard University (fall 1984 to fall 2018) was obtained from The Integrated Postsecondary Education Data System (IPEDS) database and used in this empirical study.

4.1 Predictors

Four lengths of time series ($\boldsymbol{TS} = 5, 10, 20, 30$) were compared to evaluate prediction accuracy. Using fall 2013 as the endpoint, the past five years of data from fall 2009 to fall 2013, the past ten years of data from fall 2004 to fall 2013, the past twenty years of data from fall 1994 to fall 2013, and the past thirty-years of data from fall 1984 to fall 2013 were applied as predictors to forecast the undergraduate enrollment from fall 2014 to fall 2018.

4.2 Outcome

The most recent five years of enrollment data (fall 2014 to fall 2018) were used as the true "future" data to compare with the predicted values from the ARIMA model based on the different lengths of time series.

4.3 Analysis

The AR model (1,0,0) and the differenced AR model (1,1,0) were applied in the ARIMA model in the analysis. ARIMA (4,0,0) captured the past five years' time series trajectories, ARIMA (7,0,0) captured the past ten years' time series trajectories, ARIMA (14,1,0) captured the past twenty years' time series trajectories, and ARIMA (22,1,0) captured the past thirty years' time series trajectories. The ARIMA model accurately captured the previous enrollment's changing trajectories, as shown in Figure 4. The enrollment of fall 2014 to fall 2018 was forecasted based on the past five, ten, twenty, and thirty years' trajectories, as shown in Figure 5.

4.4 Results

The differences between the forecasted enrollment and the true enrollment from fall 2014 to fall 2018 under four lengths of time series are summarized in Table 2.

The results in the empirical study were consistent with the results in the simulation study. Under the condition that the ARIMA model completely fit the past years' enrollment trajectories, the sequence of the time series at 20 had the smallest average discrepancy ($Bias3 = 375$) between the forecasted enrollment and the true enrollment from fall 2014 to fall 2018.

Three important findings came to light as part of this empirical study. The results are summarized in Table 3. First, when the ARIMA model completely fits the data, the sequence of the time series at 20 provided the highest accuracy in the enrollment forecasting. The differences between the forecasted enrollments based on the time series at 20 and the true enrollment were not significant ($t = -189.41, p = 0.97$). However, there were significant differences between the forecasted enrollment based on the time series at $5, 10,$ and 30 and the true enrollment ($p < .001$). Second, the time series at 20 had significantly higher accuracy in forecasting undergraduate enrollment than the time series at $5, 10,$ and 30 in the empirical results ($t_{TS5\ vs\ TS20} = 1839, p < 0.001; t_{TS10\ vs\ TS20} = 2369, p < 0.001; t_{TS30\ vs\ TS20} = 1529, p < 0.01$). Last, the differences among time series at $5, 10,$ and 30 were not significant ($p > .05$).

5. Conclusion and Implications

This paper proposed two comparison studies to evaluate the impact of lengths of time series on the accuracy of the ARIMA model.

A simulation study was first developed based on the average undergraduate enrollment across the top 10 HBCUs in the United States. The simulation results indicated that the shortest time series $(TS = 5)$ had the lowest forecasting accuracy, shown by the most considerable bias between the forecasted values and the simulated "true" data. When the ARIMA model accurately captured the historical changing trajectories with the lowest AIC/BIC values, the time series at $10, 20,$ and 30 had significantly higher accuracy in forecasting enrollment. Notably, the time series at 20 provided the highest accuracy in forecasting future 10 years' enrollment across 100 replications.

Similar results were also found in the empirical study. The past 35 years of undergraduate enrollment data from fall 1984 to fall 2018 at Howard University was used in the study. The most recent five years' enrollment (fall 2014 to fall 2018) was used to compare against the predicted enrollment in the ARIMA model based on different lengths of previous enrollment data. The results in the empirical study also showed that the time series at 20 provided the highest accuracy in forecasting enrollment from fall 2014 to fall 2018. The average difference between the predicted enrollment based on the time series at 20 and the true enrollment was only around 370.

In summary, based on the simulation and empirical studies, the ARIMA model is an effective and powerful statistical tool to perform time-series forecasting of enrollment in higher education. The lengths of time series lower than 10 are not sufficient to provide a relatively accurate enrollment forecasting, however. When the ARIMA model completely recovered the past changing trajectories, the ideal length of time series is 20 in enrollment forecasting. The statement that the longer time series it is, the higher forecasting accuracy is obtained in the ARIMA model was not supported in this study. One potential reason for this discrepancy is the additional time-series data brought to the higher autocorrelation and residual error in the ARIMA model. This will lead to a higher bias in the ML estimations.

The results of this paper contribute as a reference to studies in the enrollment projection and time-series forecasting. It provides a significantly practical impact on enrollment strategies, budges plans, and financial aid policies at colleges and institutions across countries. One possible way for future studies is to continue investigating the impact of lengths of time series on other time-series models, such as Holt-Winter or neural network.

References

Adrian, T., & Kurt, H. (2019). tseries: Time Series Analysis and Computational Finance. *R package version 0*, 10-47.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control.* John Wiley & Sons.

Brinkman, P. T., & McIntyre, C. (1997). Methods and Techniques of Enrollment Forecasting. *New directions for Institutional Research, 93*, 67-80. https://doi.org/10.1002/ir.9305

Chen, C. K. (2008). An integrated enrollment forecasting model. *IR Applications*, 15.

Chujai, P., Kerdprasop, N., & Kerdprasop, K. (2013, March). Time series analysis of household electric consumption with ARIMA and ARMA models. In *Proceedings of the International MultiConference of Engineers and Computer Scientists* (Vol. 1, pp. 295-300).

Colak, I., Yesilbudak, M., Genc, N., & Bayindir, R. (2015, December). Multi-period prediction of solar radiation using ARMA and ARIMA models. In *2015 IEEE 14th International Conference on Machine Learning and Applications* (ICMLA) (pp. 1045-1049). IEEE. https://doi.org/10.1109/ICMLA.2015.33

Conejo, A. J., Plazas, M. A., Espinola, R., & Molina, A. B. (2005). Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. *IEEE Transactions on Power Systems, 20*(2), 1035-1042. https://doi.org/10.1109/TPWRS.2005.846054

Contreras, J., Espinola, R., Nogales, F. J., & Conejo, A. J. (2003). ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems, 18*(3), 1014-1020. https://doi.org/10.1109/TPWRS.2002.804943

Enders, C. K. (2010). *Applied missing data analysis.* Guilford Press.

Gross, G., & Galiana, F. D. (1987). Short-term load forecasting. *Proceedings of the IEEE, 75*(12), 1558-1573. https://doi.org/10.1109/PROC.1987.13927

Keith, T. Z. (2014). *Multiple regression and beyond: An introduction to multiple regression and structural equation* *modeling.* Routledge. https://doi.org/10.4324/9781315749099

Kinney, W. R. (1978). ARIMA and regression in analytical review: An empirical test. *The Accounting Review, 53*(1), 48-60.

Marriott, J., & Newbold, P. (1998). Bayesian comparison of ARIMA and stationary ARMA models. *International Statistical Review, 66*(3), 323-336. https://doi.org/10.1111/j.1751-5823.1998.tb00376.x

Morana, C. (2001). A semiparametric approach to short-term oil price forecasting. *Energy Economics, 23*(3), 325-338. https://doi.org/10.1016/S0140-9883(00)00075-X

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research, 96*(1), 3-14. https://doi.org/10.1080/00220670209598786

RStudio Team. (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Retrieved from http://www.rstudio.com/

Valipour, M., Banihabib, M. E., & Behbahani, S. M. R. (2013). Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. *Journal of Hydrology*, *476*, 433-441. https://doi.org/10.1016/j.jhydrol.2012.11.017

Wang, C. C, Hsu, Y. S., & Liou, C. H. (2011). A comparison of ARIMA forecasting and heuristic modelling. *Applied Financial Economics*, 21-15, 1095-1102. https://doi.org/10.1080/09603107.2010.537629

Ward, J. (2007). Forecasting enrollment to achieve institutional goals. *College and University Journal*, 3-82.

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159-175. https://doi.org/10.1016/S0925-2312(01)00702-0

Appendix A

Tables

Table 1. Average bias across 100 replications between the forecasted values and the simulated "true" data

	TS 5	TS 10	TS 20	TS 30
Bias at Future Time Point 1	-42	18.55	-46.16	-81.28
Bias at Future Time Point 2	-57	-18.48	-3.69	-45.88
Bias at Future Time Point 3	-61.38	6.44	49.27	14.78
Bias at Future Time Point 4	-92.74	-11.32	7.49	-5.34
Bias at Future Time Point 5	-57.99	16.2	23.08	5.69
Bias at Future Time Point 6	-47.22	-26.5	39	4.24
Bias at Future Time Point 7	-81.05	47.65	-8.82	63.74
Bias at Future Time Point 8	-37.84	-9.31	1.4	-20.32
Bias at Future Time Point 9	-51.63	45.73	-12.13	17.73
Bias at Future Time Point 10	-22.18	-88.76	-7.89	-29.63
Average Bias	55.1	28.9	***19.9***	28.9

Table 2. Differences between the true enrollment and the forecasted enrollment from Fall 2014 to Fall 2018

	True Enrollment	TS 5	Bias1	TS 10	Bias2	TS 20	Bias3	TS 30	Bias4
Fall 2014	7013	7317	304	8827	1814	6773	-240	7243	230
Fall 2015	6883	8910	2027	8950	2067	6940	57	7692	809
Fall 2016	5899	7980	2081	9017	3118	6456	557	7854	1955
Fall 2017	6354	8546	2192	8584	2230	6031	-323	7897	1543
Fall 2018	6243	7885	1642	7913	1670	5544	-699	8406	2163
Average Bias			1649.2		2179.8		***375.2***		1340

Table 3. The t-test of the ARIMA model based on the true enrollment and four lengths of time series forecasting

	True	TS 5	TS 10	TS 20	TS 30
True	0				
TS 5	t = 1649, ***	0			
TS 10	t = 2180, ***	t = 530, ns	0		
TS 20	***t = 189, ns***	t = 1839, ***	t = 2369, ***	0	
TS 30	t = 1340, ***	t = 309, ns	t = 839, ns	t = 1529, ***	0

*Note. *** p < .001; ns p > .05*

Appendix B

Figures

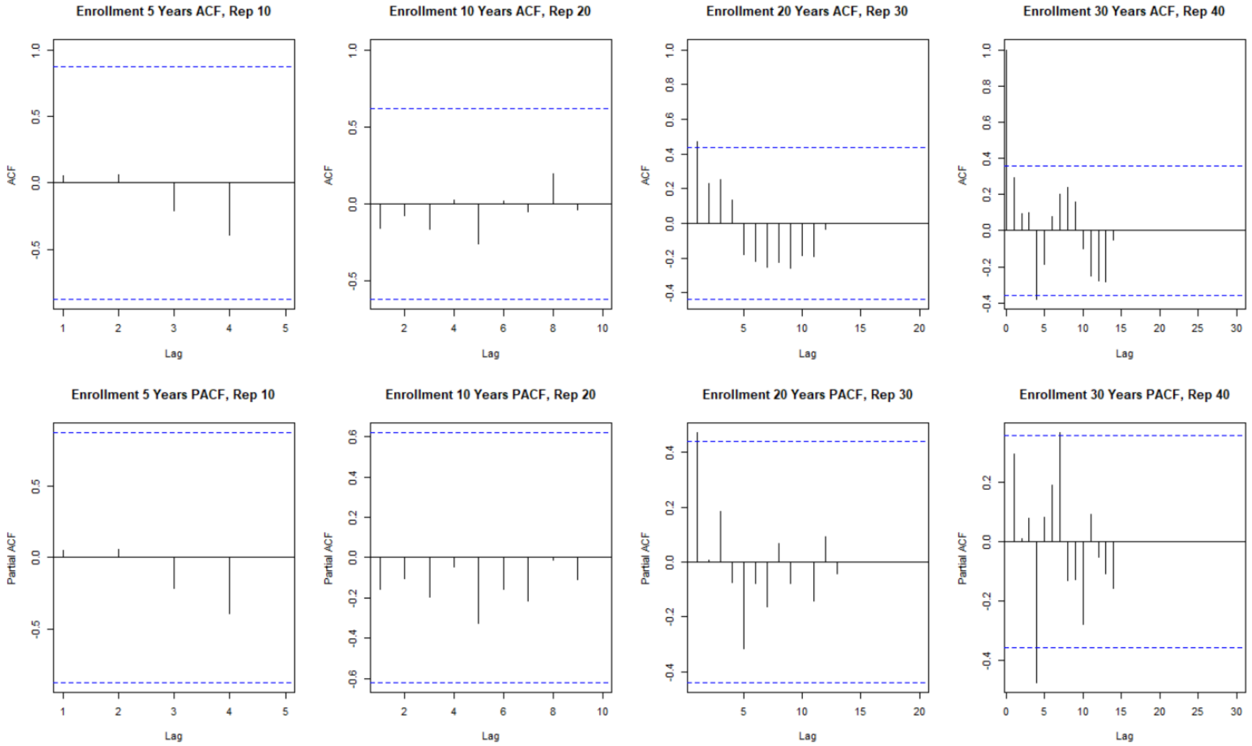

Figure 1. ACF and PACF plots examples

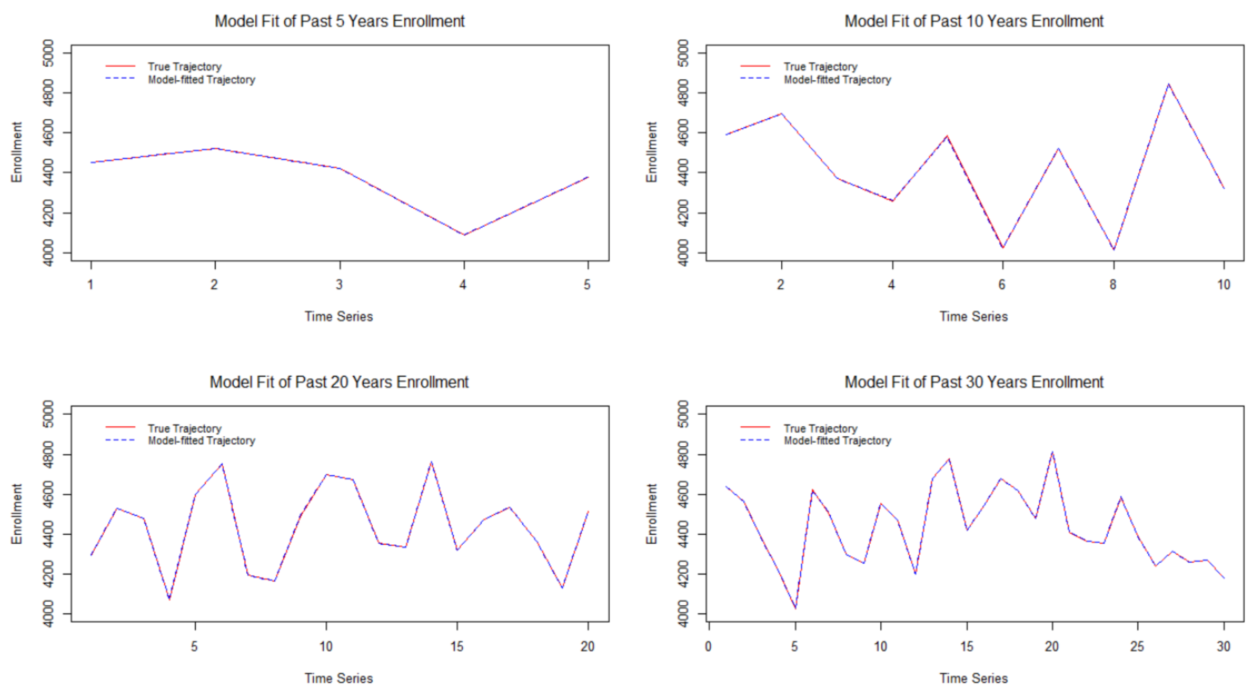

Figure 2. Model Fit of Simulated Data

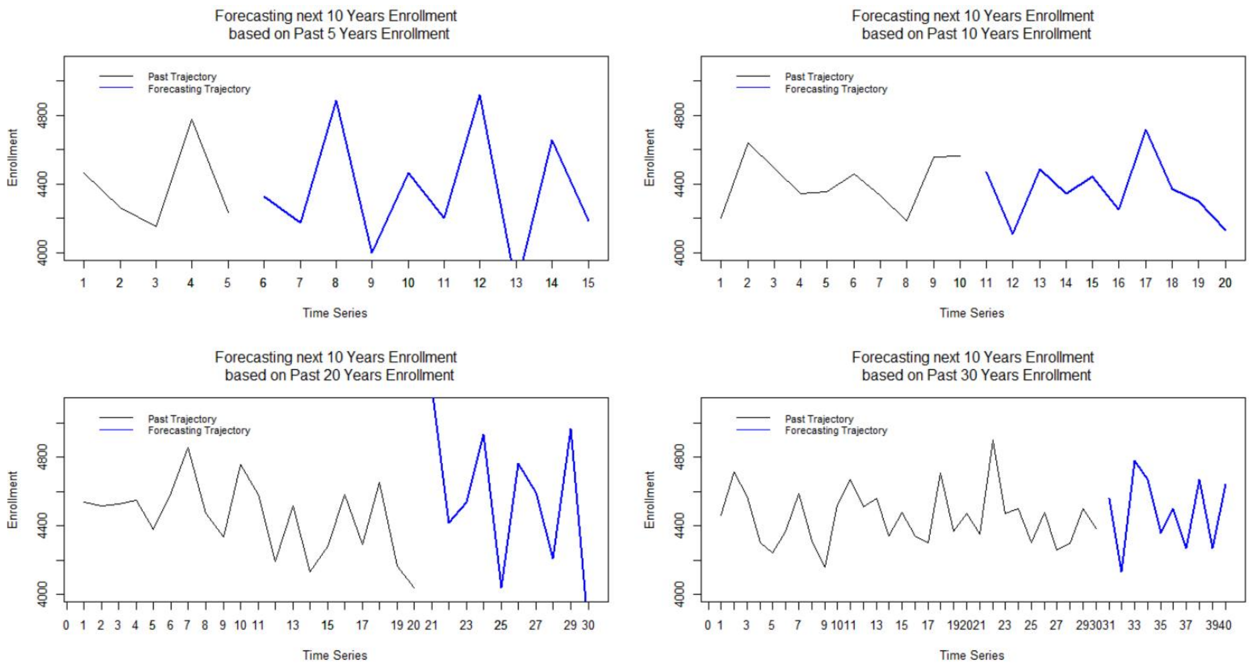

Figure 3. Forecasting Future Changing Patterns

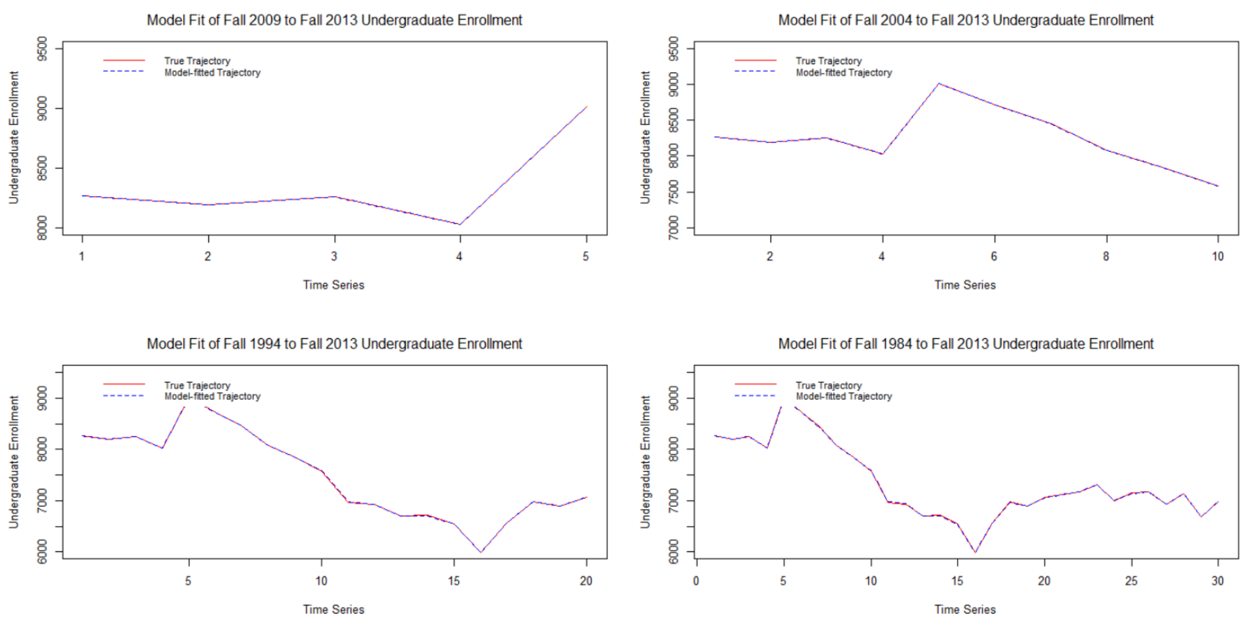

Figure 4. Model Fit of Empirical Data

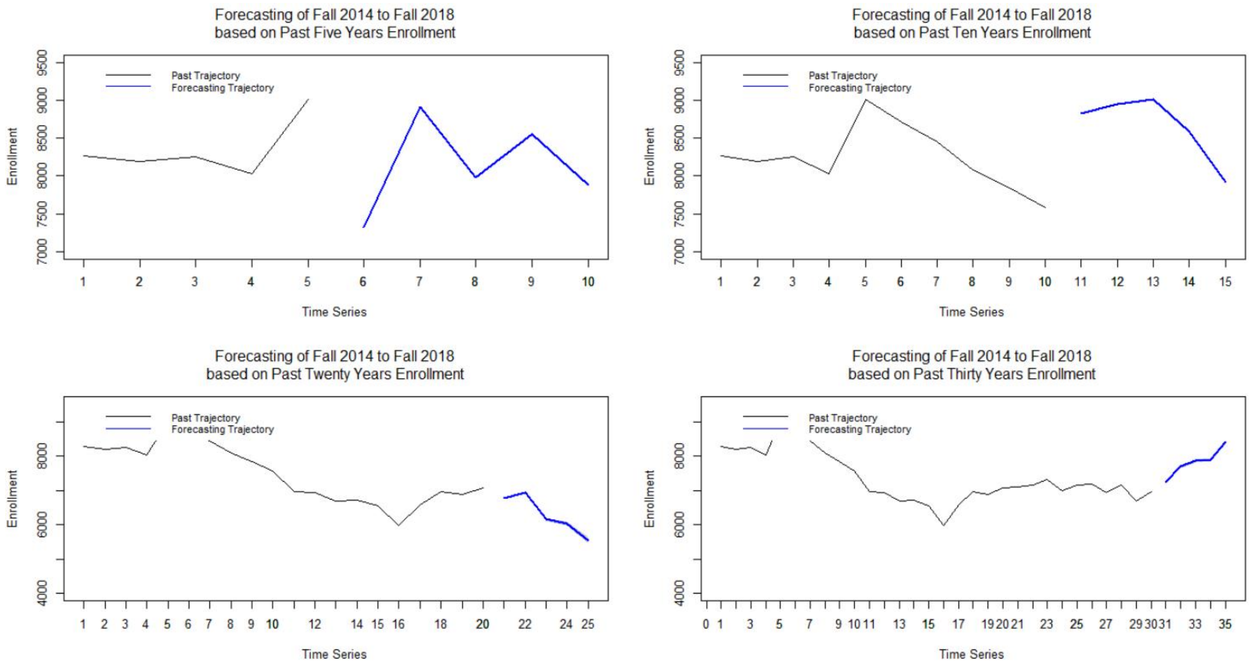

Figure 5. Forecasting of Enrollment from Fall 2014 to Fall 2018