Journal of Applied Data Science
Vol. 2, No. 2, May 2021, pp. 01-13

Modeling and Forecasting Long-Term Records of Mean Sea Level at Grand Isle, Louisiana: SARIMA, NARNN, and Mixed SARIMA-NARNN Models

Yeong Nain Chi [1,*]

[1] Department of Agriculture, Food and Resource Sciences, University of Maryland Eastern Shore, Princess Anne, MD 21853, USA
[1] ychi@umes.edu*
* corresponding author

(Received March 3, 2021 Revised April 4, 2021 Accepted April 13, 2021, Available online May 1, 2021)

Abstract

This study tried to demonstrate the role of time series models in modeling and forecasting process using long-term records of monthly mean sea level from January 1978 to October 2020 at Grand Isle, Louisiana. Following the Box–Jenkins methodology, the ARIMA$(1,1,1)(2,0,0)_{12}$ with drift model was selected to be the best fit model for the time series, according to its lowest AIC value. Using the LM algorithm, the results revealed that the NARNN model with 9 neurons in the hidden layer and 6 time delays provided the best performance in the nonlinear autoregressive neural network models at its smaller MSE value. The Mixed model, a combination of the SARIMA and NARNN models has both linear and nonlinear modelling capabilities can be a better choice for modelling the time series. The comparative results revealed that the Mixed-LM model with 9 neurons in the hidden layer and 3 time delays yielded higher accuracy than the NARNN-LM model with 9 neurons in the hidden layer and 6 time delays, and the ARIMA$(1,1,1)(2,0,0)_{12}$ with drift model, according to its lowest MSE in this study. Thus, this study may provide an integrated modelling approach as a decision-making supportive method for formulating local mean sea level forecast in advance. Understanding past sea level is important for the analysis of current and future sea level changes. In order to sustain these observations, research programs utilizing the resulting data should be able to improve significantly our understanding and narrow projections of future sea level rise and variability.

Keywords: Mean Sea Level; Time Series; Modeling; Forecasting; SARIMA; NARNN; Mixed SARIMA-NARNN Model; Levenberg-Marquardt Algorithm; Grand Isle, Louisiana.

1. Introduction

Climate change has been addressed intensively. At the same time, the issues of sea level rise at an increasing rate has been discussed comprehensively as well. There are two major factors related to climate change caused sea level rise globally: the added water from ice melting (ice sheets and glaciers) into the ocean, and the thermal expansion of warming waters. Locally, the amount and speed of sea level rise varies by location, particularly, the slowing Gulf Stream and sinking land affect some areas at varying rates in the United States (https://sealevelrise.org/). National Oceanic and Atmospheric Administration (NOAA) 2019 Global Climate Annual Report summarized that the global annual temperature has increased at an average rate of 0.07°C (0.13°F) per decade since 1880. This rate (+0.18°C / +0.32°F) of increase has doubled since 1981 (https://www.ncdc.noaa.gov/sotc/global/201913).

According to NOAA Climate.gov, global average sea level has risen about 8-9 inches (21-24 cm) since 1880. The Intergovernmental Panel on Climate Change (IPCC) [18] estimated that the sea level has risen by 26-55 cm (10-22 inches) with a 67% confidence interval. If emissions remain very high, the IPCC projected sea level could rise by 52-98 cm (20-39 inches). In its Fourth National Climate Assessment Report [27] the U.S. Global Change Research Program (USGCRP) estimated that sea level has risen by about 7-8 inches (about 16-21 cm) since 1900, with about 3 of those inches (about 7 cm) occurring since 1993 (very high confidence). Relative to the year 2000, sea level was very likely to rise by 1.0-4.3 feet (30-130 cm) in 2100, and 0.3-0.6 feet (9-18 cm) by 2030.

There were many studies pointed out that sea level is rising at an increasing rate [10] [11] [12] [8] [13] [9] [17] [21] [16]. Thus, understanding past sea level is important for the analysis of current and future sea level changes. Modeling sea level changes and understanding its causes has considerably improved in the recent years, essentially because new in situ and remote sensing observations have become available [14] [28] [3] [25]. Despite the

importance of sea level rise and its consequences, there is a lack of studies in the technical literature available on forecasting schemes at the local consideration.

Recently, there is increasing interest in using neural networks to model and forecast time series [30], particularly in the sea level rise issues [23] [6] [29] [7]. Hence, the primary purpose of this study was to demonstrate the role of time series model in predicting process and to pursue analysis of time series data using long-term records of monthly mean sea level from January 1978 to October 2020 at Grand Isle, Louisiana.

2. Materials

2.1. Study Site

Grand Isle is a town at latitude 29.2366 and longitude -89.9873 in Jefferson Parish, Louisiana, located on a barrier island at the mouth of Barataria Bay where it meets the Gulf of Mexico (Figure 1). According to the United States Census Bureau, the town covered a total area of 7.8 square miles (20 km^2), of which 6.1 square miles (16 km^2) is land and 1.7 square miles (4.4 km^2) is water. Grand Isle has a humid subtropical climate bordering on a tropical monsoon climate, with mild winters and long, hot, humid summers. Grand Isle has been repeatedly pummeled by hurricanes through its history. On average, Grand Isle has been affected by tropical storms or hurricanes every 2.68 years since 1877, with direct hits on average every 7.88 years. Hurricane Katrina, for example, pounded Grand Isle for two days, August 28–29, 2005, destroying or damaging homes and camps along the entire island. Katrina's surge reached 5 ft (1.5 m) at Grand Isle.

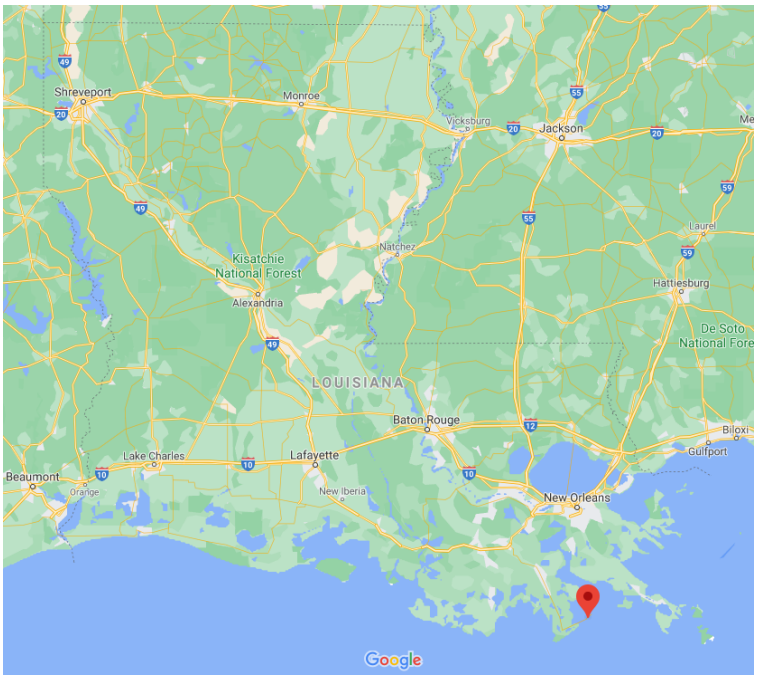

Fig. 1. Grand Isle, Louisiana, USA. Source: Google Map (adapted from: https://www.google.com/maps/place/Grand+Isle,+LA/**)**

2.2. Data Source

The long-term records of monthly mean sea level from January 1978 to October 2020 at Grand Isle, Louisiana, used for this study is available to the public from NOAA Tides and Currents (https://tidesandcurrents.noaa.gov/). According to NOAA Tides and Currents, the term mean sea level can refer to a tidal datum, which is locally-derived based on observations at a tide station, and is typically computed over a 19-year period, known as the National Tidal Datum Epoch (NTDE). Mean sea level as a tidal datum is computed as a mean of hourly water level heights observed over 19-years. Monthly means generated in the datum calculation process, which is used to generate the relative local sea level trends observed at a tide station.

ISSN 2579-7069

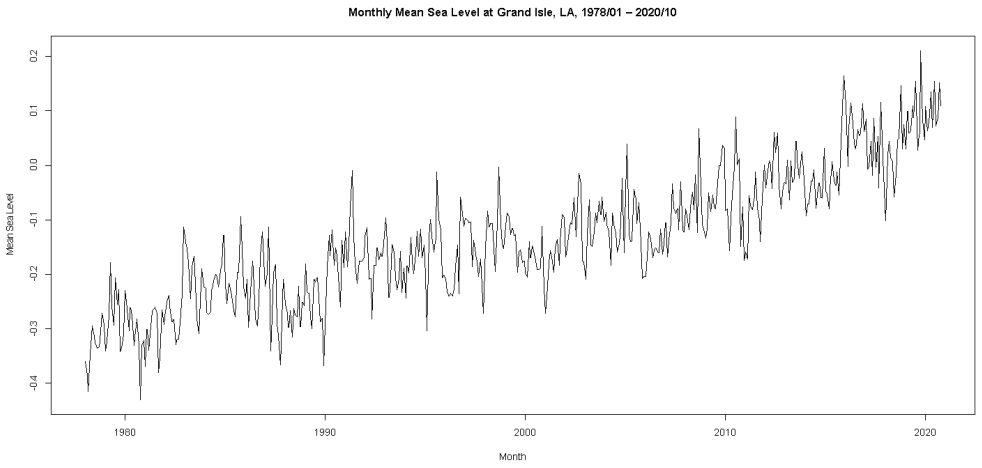

Fig. 2. Time Series Plot of Monthly Mean Sea Level at Grand Isle, Louisiana, January 1978 ~ October 2020 (Source: own work)

3. Methods

3.1. Seasonal ARIMA (SARIMA) Model

A time series is a set of observations, each one being recorded at a specific time t. The sequence of random variables $\{y_t: t = 1, 2, \cdots, T\}$ is called a stochastic process and serves as a model for an observed time series. Statistically, the ARIMA(p,d,q) model can be expressed in backshift notation as:

$$\phi_p(B)(1 - B)^d y_t = c + \theta_q(B)e_t \quad (1)$$

where p = the order of the autoregressive process, d = the number of differences required to make the time series stationary, q = the order of the moving average process, $\phi_p(B) = (1 - \phi_1 B - \ldots - \phi_p B^p) = (1 - \Sigma^p{}_{i=1}\ \phi_i B^i)$, $\theta_q(B) = (1 - \theta_1 B - \ldots - \theta_q B^q) = (1 - \Sigma^q{}_{j=1}\ \theta_j B^j)$, c is a constant, and e_t is the error term. The purpose of each of these parts is to make the model better fit to predict future points in the time series [24].

The ARIMA(p,d,q)(P,D,Q)$_S$ model is used to represent the SARIMA model that can be written in backshift notation as:

$$\phi_p(B)\Phi_p(B^m)(1 - B)^d(1 - B^m)^D y_t = c + \theta_q(B)\Theta_Q(B^m)e_t \quad (2)$$

where P = the order of the seasonal autoregressive process, D = the number of seasonal differences applied to the time series, Q = the order of the seasonal moving average process, S = the seasonality of the model (i.e., the number of time steps for a single seasonal period), $\Phi_P(B^m) = (1 - \Phi_1 B^m - \ldots - \Phi_P B^{mP}) = (1 - \Sigma^P{}_{i=1}\ \Phi_i B^{mi})$, and $\Theta_Q(B^m) = (1 - \Theta_1 B^m - \ldots - \Theta_Q B^{mQ}) = (1 - \Sigma^Q{}_{j=1}\ \Theta_j B^j)$.

In time series forecasting, the Box-Jenkins methodology [4] refers to a systematic method of identifying, estimating, checking, and forecasting ARIMA models [5], that can be applied to find the best fit of the time series. The Box-Jenkins methodology also can be used as the process to forecast the SARIMA model in this study based on its autocorrelation function (ACF) and partial autocorrelation function (PACF) as a means of determining the stationarity of the univariate time series and the lag lengths. Thus, the Box-Jenkins methodology starts with the assumption that the time series should be as on stationary status. If the time series is not stationary, it needs to be stationarized through differencing.

Empirically, plots and summary statistics can be used to identify trends and autoregressive elements to get an idea of the amount of differencing and the size of the lag that will be required for model identification. In order to figure out good parameters for the model, Akaike's Information Criterion (AIC) or Bayesian Information Criterion (BIC) can be used to determine the orders of a SARIMA model that is obtained by minimizing the AIC or BIC value. In the

diagnostic checking step, plots and statistical tests of the residual errors can be used to determine the model fitting, to evaluate the fitting model in the context of the available data, and to check where the model may be improved.

3.2. Nonlinear Autoregressive Neural Network (NARNN) Model

The idea behind the autoregressive (AR) process is to explain the present value of the time series, y_t, by a function of p past values, (y_{t-1}, y_{t-2}, ⋯, y_{t-p}). Thus, the AR process of order p, AR(p), is defined by the equation:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t = \sum_{i=1}^{p} \phi_i y_{t-i} + e_t \quad (3)$$

where $\phi = (\phi_1, \phi_2, \cdots, \phi_p)$ is the vector of model coefficients for the autoregressive process, and e_t is white noise, i.e., $e_t \sim N(0, \sigma^2)$. The NARNN is a natural generalization of the classic linear AR(p) process. The NARNN of order p can be expressed as:

$$y_t = \Phi(y_{t-1}, y_{t-2}, \cdots, y_{t-p}, w) + \varepsilon_t \quad (4)$$

where $\Phi(\cdot)$ is an unknown function determined by the neural network structure and connection weights, w is a vector of all parameters (weights), and ε_t is the error term. Thus, it performs a nonlinear functional mapping from the past observations, (y_{t-1}, y_{t-2}, ⋯, y_{t-p}), to the future value, y_t, which is equivalent to a nonlinear autoregressive model [30]. With the time series data, lagged values of the time series can be used as inputs to a neural network, so-called this the NARNN model. Mathematically, the NARNN model [2] can be written by the equation of the form as:

$$y_t = a_0 + \sum_{j=1}^{k} w_j \, \Phi(b_{0j} + \sum_{i=1}^{d} w_{ij} y_{t-i}) + \varepsilon_t \quad (5)$$

where d is the number of input units, k is the number of hidden units, a_0 is the constant corresponding to the output unit, b_{0j} is the constant corresponding to the hidden unit j, w_j is the weight of the connection between the hidden unit j and the output unit, w_{ij} is the parameter corresponding to the weight of the connection between the input unit i and the hidden unit j, and $\Phi(\cdot)$ is a nonlinear function, so-called this the transfer (activation) function. The logistic function (i.e., sigmoid) is commonly used as the hidden layer transfer function, that is, $\Phi(y) = 1 / (1 + \exp(-y))$.

The most common learning rules for the NARNN model are the Levenberg-Marquardt, Bayesian Regularization, and Scaled Conjugate Gradient training algorithms. In this study, the Levenberg-Marquardt (LM) algorithm was considered, because it works without computing the exact Hessian matrix. Instead, it works with the gradient vector and the Jacobian matrix, therefore increasing the training speed and has stable convergence [15].

3.3. Mixed SARIMA-NARNN (Mixed) Model

The SARIMA and NARNN models are good at modelling linear and nonlinear problems for the time series, respectively. However, using the Mixed model, a combination of the SARIMA and NARNN models has both linear and nonlinear modelling capabilities, can be a better choice for modelling the time series. Assuming that an unknown function can be used to demonstrate the relationship between linear and nonlinear components in the time series, which can be illustrated as follows:

$$y_t = f(L_t, N_t) \quad (6)$$

where linear component is represented by L_t, and nonlinear component is shown by N_t. Assuming that the linear and nonlinear components in the time series have simply additive relationships. Zhang [30] states that the time series can be considered as a combination of a linear and nonlinear components as follows:

$$y_t = L_t + N_t \quad (7)$$

These two components should be estimated from the time series. First, the linear component will be modelled by the SARIMA model in this study. Then, the residuals from the SARIMA model will have only the nonlinear relationship, which can be obtained by taking difference of observed values and predicted values as follows:

ISSN 2579-7069

$$e_t = y_t - \hat{L}_t \quad (8)$$

where e_t is the residual of the linear model at time t, and $\hat{L}_t$ is the predicted value for time t. To find the nonlinear relationship, residuals can be modelled by the NARNN model in this study as follows:

$$\hat{N}_t = e_t = f(e_{t-1}, e_{t-2}, \cdots, e_{t-n}) + \varepsilon_t \quad (9)$$

where f is the transformation function modelled by the NARNN model, and ε_t is the random error. The forecast from the SARIMA and NARNN models are combined to obtain the forecast of the time series $\hat{y}_t$ which is denoted by

$$\hat{y}_t = \hat{L}_t + \hat{N}_t. \quad (10)$$

4. Results

4.1. Seasonal ARIMA (SARIMA) Model

R 4.0.2 for Windows is an open source for statistical computing and graphics supported by the R Foundation for Statistical Computing was used as the tool to model and forecast monthly mean sea level from January 1978 to October 2020 at Grand Isle, Louisiana to achieve the purpose of this study. The function "decompose()" in R can be applied to separate the seasonal component, trend component and irregular components of a seasonal time series. The estimated trend component showed a steady increase over time, and the estimated seasonal component definitely displayed seasonality, with a pattern recurrence occurring once every 12 months (yearly) (Figure 3).

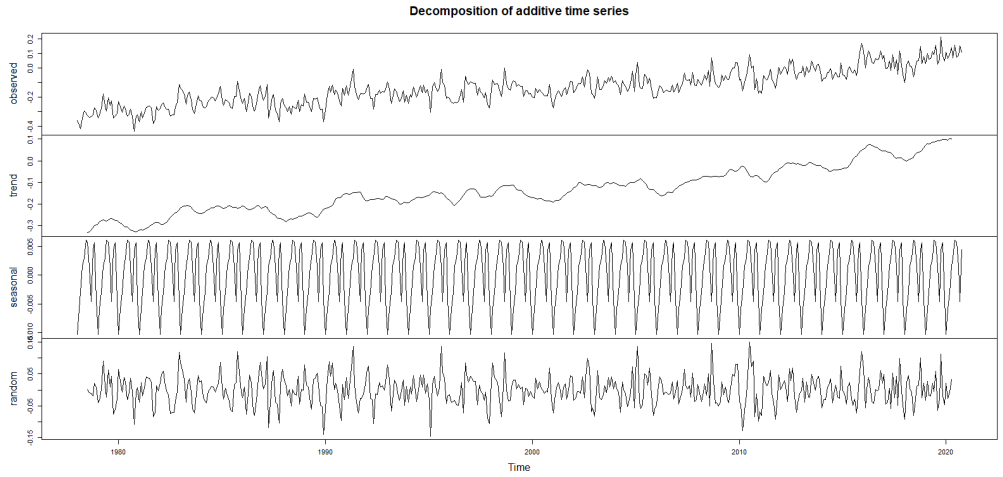

Fig. 3. Decomposition of Monthly Mean Sea Level at Grand Isle, Louisiana, January 1978 ~ October 2020 (Source: own work)

Seasonal adjustment is the removal of seasonal effects that are not explainable by the dynamics of trends or cycles from a time series to reveal certain non-seasonal features. This can be done by subtracting the estimated seasonal component from the original time series. After removed the seasonal variation, the seasonally adjusted time series only contained the trend component and an irregular component. The ACF of the time series, seasonal adjusted monthly mean sea level from January 1978 to October 2020 at Grand Isle, Louisiana, showed strong positive statistically significant correlations at up to 28 lags that never decay to zero. This also illustrated by the single spike at the first lag followed by small apparently random values after the first lag for the PACF.

In terms of the non-stationary time series, differencing can be used to transform a non-stationary time series into a stationary one. When both trend and seasonality are present, a non-seasonal first difference and a seasonal difference may need to apply in advance. Notice that the graph of the first difference of the time series looked approximately stationary. According to the Augmented Dickey-Fuller Test, Dickey-Fuller = -11.944 with lag order = 7 and the p-value of the test was smaller than 0.01. It rejected the null hypothesis that is non-stationary, and suggested that the first difference of the time series was stationary. The ACF of first difference (Figure 4) showed a significant

positive spike at the first lag followed by correlations that were statistically significant. The corresponding PACF of first difference (Figure 5) showed most likely a gradual decrease after the first few lags.

Seasonal differencing is defined as a difference between a value and a value with lag that is a multiple of seasonality (S). In this case, S = 12 (months per year) is the span of the periodic seasonal behavior. Figure 5 showed the graph of the 12th difference of the time series, which looked approximately stationary. At the same time, the test statistic of the Augmented Dickey-Fuller Test was Dickey-Fuller = -37.923 with lag order = 7 and the p-value of the test was smaller than 0.01. It rejected the null hypothesis that is non-stationary, and suggested that the 12th difference of the time series was stationary. The ACF of 12th difference (Figure 6) most likely a steady decay after the first few lags and bounce around between being positive and negative statistically significant. Meanwhile, the PACF of 12th difference (Figure 7) mostly looks like a steady negative decay in the partial correlations toward zero. This is consistent with the general recommendation that if the autocorrelation at the first seasonal lag is positive we should use an autoregressive (AR) model vs. a moving average (MA) model.

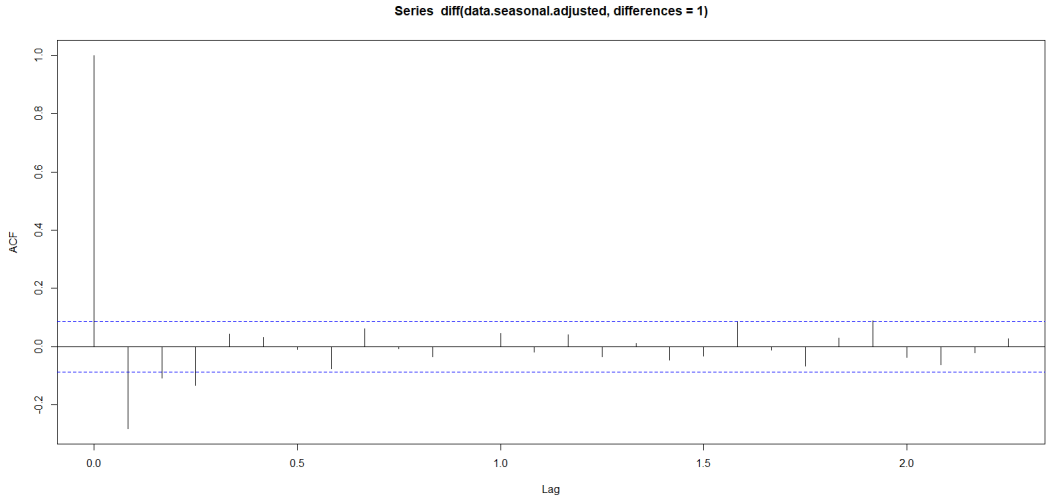

Fig. 4. ACF Plot of First Difference of Seasonal Adjusted Monthly Mean Sea Level at Grand Isle, Louisiana, January 1978 ~ October 2020 (Source: own work)

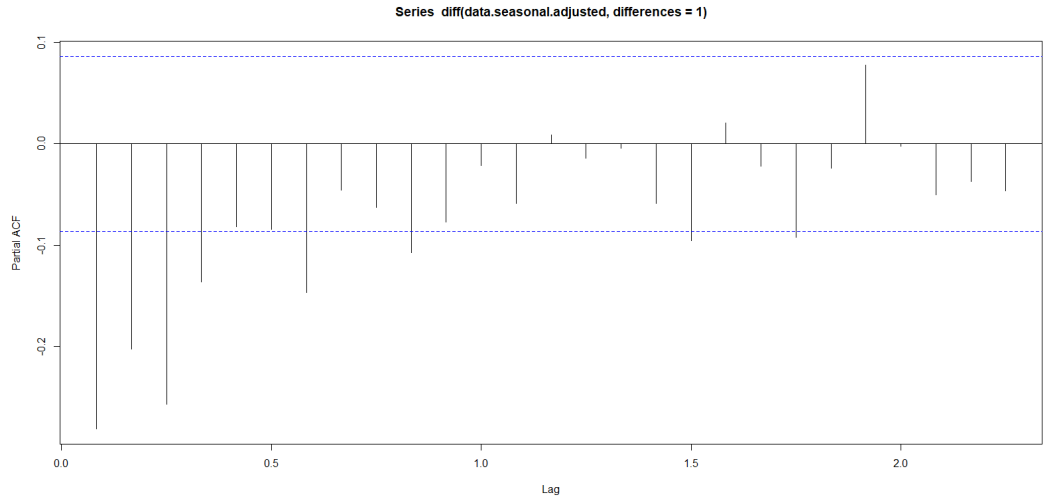

Fig. 5. PACF Plot of First Difference of Seasonal Adjusted Monthly Mean Sea Level at Grand Isle, Louisiana, January 1978 ~ October 2020 (Source: own work)

ISSN 2579-7069

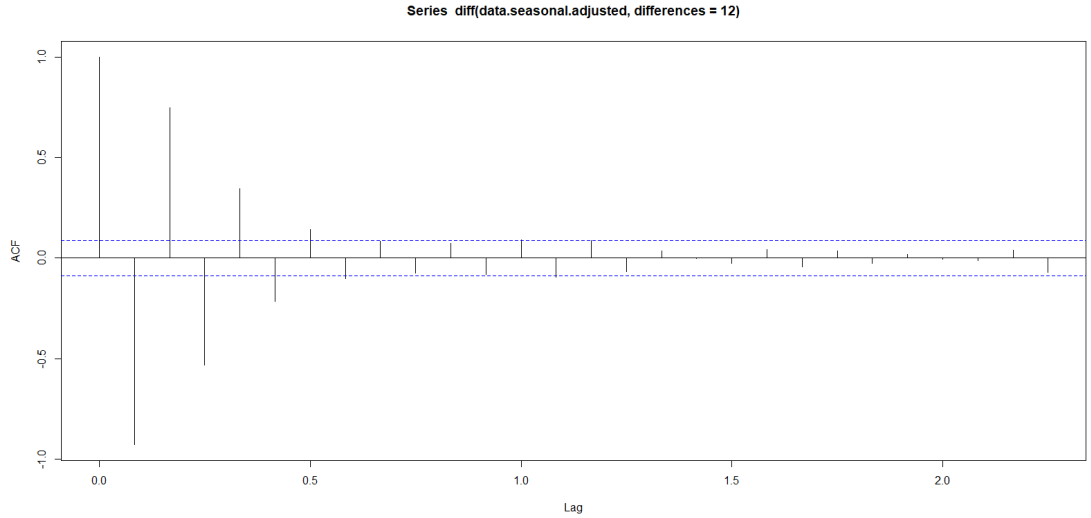

Fig. 6. ACF Plot of 12th Difference of Seasonal Adjusted Monthly Mean Sea Level at Grand Isle, Louisiana, January 1978 ~ October 2020 (Source: own work)

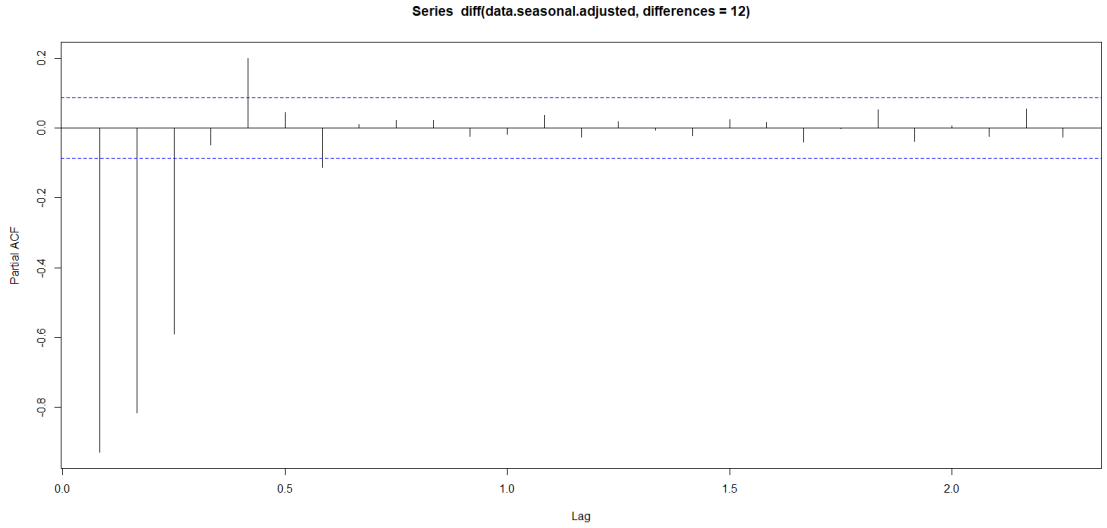

Fig. 7. PACF Plot of 12th Difference of Seasonal Adjusted Monthly Mean Sea Level at Grand Isle, Louisiana, January 1978 ~ October 2020 (Source: own work)

The auto.arima() function from the "forecast" package in R 4.0.2 for Windows was employed to identify both the structure of the time series (stationary or not) and type (seasonal or not), and sets the model's parameters, that takes into account the AIC, AICc or BIC values generated to determine the best fit SARIMA model.

Consequently, the ARIMA(1,1,1)(2,0,0)$_{12}$ with drift model was selected to be the best fit model for the time series in this study, according to the lowest AIC value (-1612.92). The ARIMA(1,1,1)(2,0,0)$_{12}$ with drift model would yield the following forecasting equation:

$$(1 - \phi_1 B)(1 - B)(1 - \Phi_1 B^{12} - \Phi_2 B^{24})\hat{Y}_t = c + (1 - \theta_1 B) \quad (11)$$

The ARIMA(1,1,1)(2,0,0)$_{12}$ with drift model for the time series, seasonal adjusted monthly mean sea level from January 1978 to October 2020 at Grand Isle, Louisiana, can be expressed as follows:

$$(1 - 0.4627B)(1 - B)(1 - 0.0542B^{12} + 0.0647B^{24})\hat{Y}_t = 0.0008 + (1 + 0.9671B) \quad (12)$$

ISSN 2579-7069

A common task when building a forecasting model is to check that the residuals satisfy some assumptions that they are uncorrelated, normally distributed, etc. The top figure of Figure 8 showed that the residuals from the ARIMA(1,1,1)(2,0,0)$_{12}$ with drift model did not violate the assumption of constant location and scale. The bottom right figure of Figure 8 also showed that the residual histogram did not reveal a time series deviation from normality in this case. The ACF plot of the residuals (the bottom left figure of Figure 8) illustrated that all sample autocorrelations were within the threshold limits, indicating that the residuals appeared to be random.

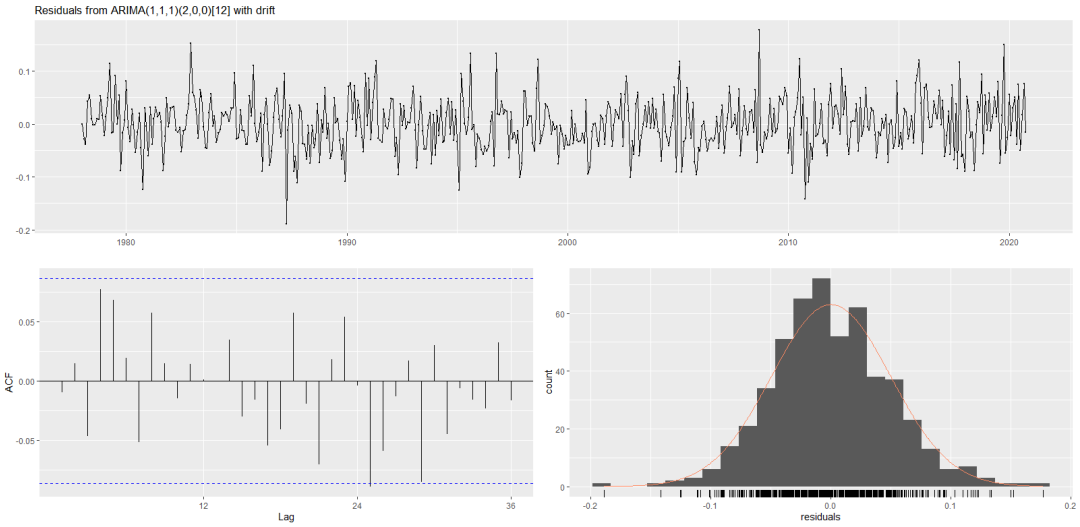

Fig. 8. Residuals of ACF and PACF (Source: own work)

In this study, the test statistic of the Ljung-Box Q-test [22] was Q = 20.637 with 19 degrees of freedom and the p-value of the test was 0.3572 (model degrees of freedom: 5, total lags used: 24), indicating that the residuals were random and that the model provided an adequate fit to the time series data. This, combined with the Ljung-Box Q-test statistic, suggested that the ARIMA(1,1,1)(2,0,0)$_{12}$ with drift model appropriately modeled the dynamics for this time series. The following figure (Figure 9) illustrated that the black line represented the visuals of monthly mean sea level dataset without forecasting and the red line represented the visuals of monthly mean sea level dataset with forecasted values. Forecasting process with the ARIMA(1,1,1)(2,0,0)$_{12}$ with drift model indicated a good fit of the model for forecasting at a constant increasing rate in the short-term.

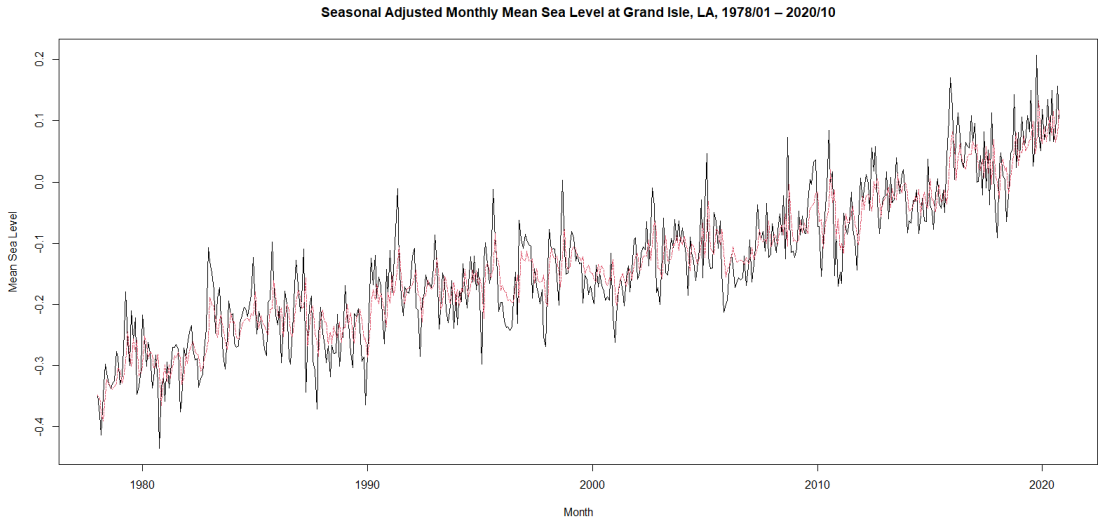

Figure 9. Observed and Forecasted Monthly Mean Sea Level at Grand Isle, Louisiana (Source: own work)

4.2. Nonlinear Autoregressive Neural Network (NARNN) Model

In MATLAB, the NARNN model applied to time series prediction using its past values of a univariate time series can be expressed as follows:

$$y(t) = \Phi(y(t-1), y(t-2), \cdots, y(t-d)) + e(t) \quad (13)$$

where y(t) is the time series value at time t, d is the time delay, and e(t) is the error of the approximation of the time series at time t. The function $\Phi(\cdot)$ is an unknown nonlinear function, and the training of the neural network aims to approximate the function by means of the optimization of the network weights and neuron bias. This tends to minimize the sum of the squared differences between the observed (y_i) and predicted ($\hat{y}_i$) values (i.e., MSE = (1/n) $\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$) [1].

In this study, the NARNN model was applied to model and forecast the time series, monthly mean sea level from January 1978 to October 2020 at Grand Isle, Louisiana. Furthermore, the logistic sigmoid and linear transfer functions at the hidden and output layers were used respectively. The extracted features were trained using the LM training algorithm for the target time series in the MATLAB (2019a) Neural Network Toolbox: 514 timesteps of one element, monthly mean sea level from January 1978 to October 2020 at Grand Isle, Louisiana. Three kinds of target timesteps were set aside for the training, validation and testing phases in this case study. The division of the time series in this analytical work was 70% for the training, 15% for the validation, and 15% for the testing. Randomly, 514 data samples were divided into 360 data for the training, 77 data for the validation, and 77 data for the testing.

The development of the optimal architecture for the NARNN model requires determination of time delays, the number of hidden neurons, and an efficient training algorithm. The optimum number of time delays and hidden neurons were obtained through a trial and error procedure. The prediction performance of the models was evaluated by its MSE. The error analysis showed that the NARNN model with 9 neurons in the hidden layer and 6 time delays provided the best performance (MSE = 2.32078e-3) using the LM algorithm (NARNN-LM). The results revealed the training progress using the LM algorithm stopped when the validation error increased for six iterations with Performance = 0.00213, Gradient = 0.000585, and Mu = 1.00e-05 at epoch 9. In terms of the processing time, the LM algorithm took 0:00:00 during the training process.

As illustrated in the performance plot, the best performance for the validation phase was 0.0028683 at epoch 3 for the NARNN-LM model. The results showed a good network performance because the validation and testing errors had similar characteristics, and did not appear that any significant overfitting has occurred. As shown in the regression plots, the regression R value for the training phase was 0.90731, for the validation phase was 0.87905, for the testing phase was 0.88299, and for the all samples was 0.89865, respectively, indicated good predictive abilities of the NARNN-LM model for new datasets.

The dynamic network time-series response plots were displayed in Figure 10 for the NARNN-LM model, showed that the outputs were distributed evenly on both sides of the response curve, and the errors versus time were small in the training, validation, and testing phases, indicated that the NARNN-LM model was able to predict the time series over the simulation period efficiently. The error autocorrelation function describes how the prediction errors are related in time. For a perfect prediction model, there should only be one nonzero value of the autocorrelation function, and it should occur at zero lag (i.e., MSE). This would mean that the prediction errors are completely uncorrelated with each other (white noise). The correlations for the NARNN-LM model (Figure 11) except for the one at zero lag, almost all fell approximately within the 95% confidence limits around zero, so the model seemed to be adequate.

ISSN 2579-7069

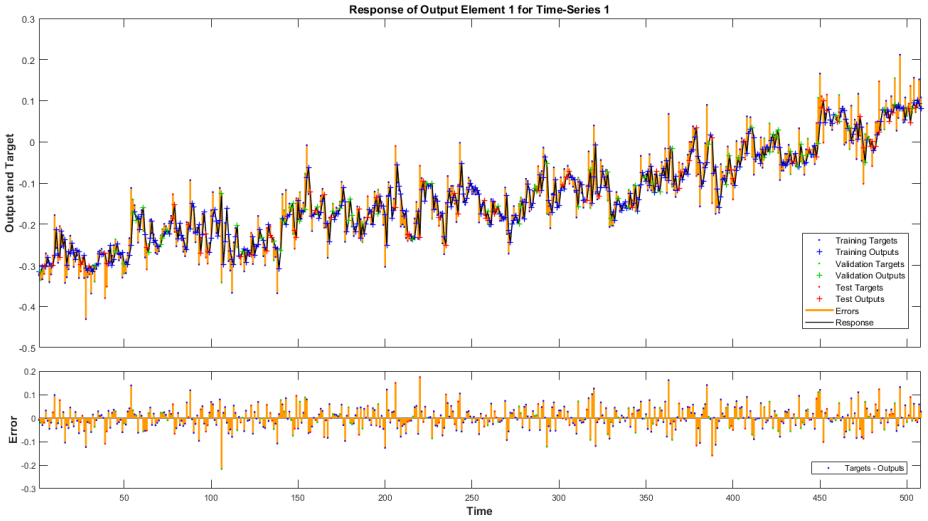

Fig. 10. Network Time-Series Response of the NARNN-LM Model (Source: own work)

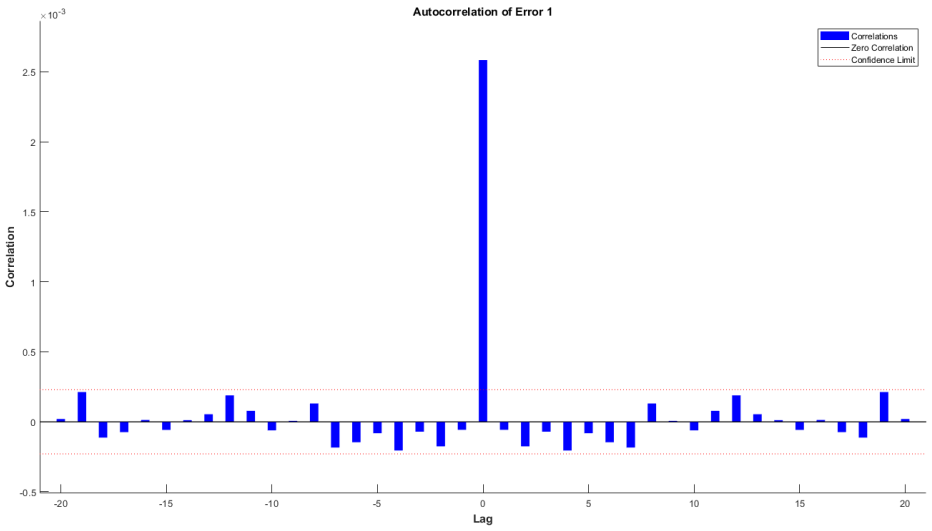

Fig. 11. Error Autocorrelation of the NARNN-LM Model (Source: own work)

4.3. Mixed SARIMA-NARNN (Mixed) Model

In MATLAB, the Mixed model applied to time series prediction using its past residuals from the SARIMA model can be expressed as follows:

$$e(t) = \Phi(e(t-1), e(t-2), \cdots, e(t-d)) + \varepsilon(t) \quad (14)$$

where e(t) is the residual of the time series at time t, d is the time delay, and ε(t) is the error term. This equation describes how the Mixed model is used to predict the future residual of a time series, e(t), using the past residuals of the time series, (e(t-1), e(t-2), ⋯, e(t-d)).

Similarly, the development of the optimal architecture for the Mixed model requires determination of time delays, the number of hidden neurons, and an efficient training algorithm. According to the results of the error analysis, it showed that the Mixed model with 9 neurons in the hidden layer and 3 time delays also provided the best performance (MSE = 2.29385e-3) with the LM algorithm (Mixed-LM). At the same time, the training progress using

ISSN 2579-7069

the LM algorithm for the Mixed-LM model stopped when the validation error increased for six iterations with Performance = 0.00226, Gradient = 0.000875, and Mu = 1.00e-06 at epoch 9.

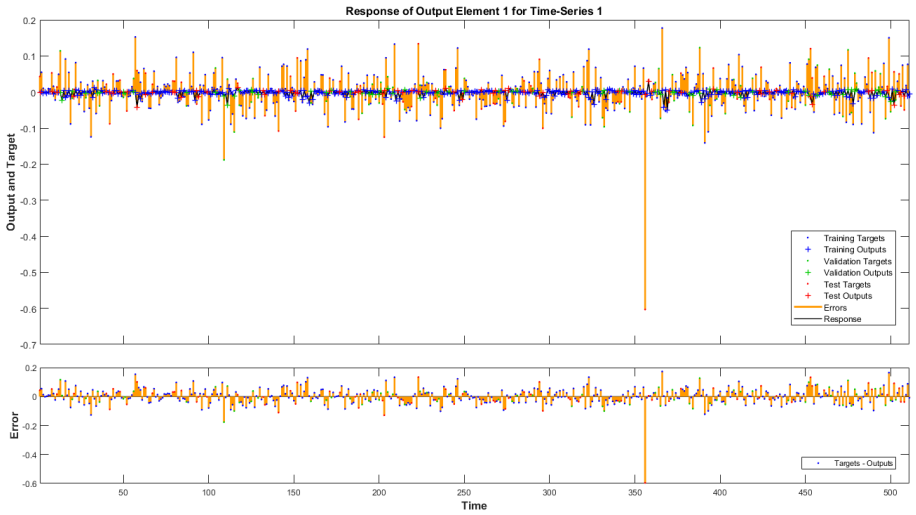

Fig. 12. Network Time-Series Response of the Mixed-LM Model (Source: own work)

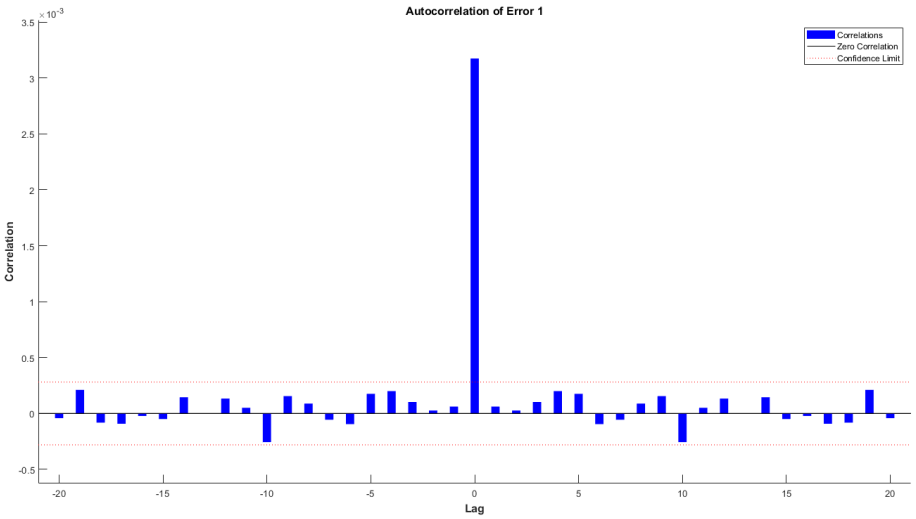

Fig. 13. Error Autocorrelation of the Mixed-LM Model (Source: own work)

As illustrated in the performance plot, the best performance for the validation phase was 0.0028901 at epoch 3 for the Mixed-LM model. The results also showed a good network performance because the validation and testing errors have similar characteristics, and did not appear that any significant overfitting has occurred. The dynamic network time-series response plots were displayed in Figure 12 for the Mixed-LM model, showed that the outputs were distributed evenly on both sides of the response curve, and the errors versus time were small in the training, validation, and testing phases. The results indicated that the Mixed-LM model was able to predict the time series over the simulation period efficiently as well. For the Mixed-LM model, the correlations except for the one at zero lag, all fell approximately within the 95% confidence limits around zero, so the model was to be adequate (Figure 13).

ISSN 2579-7069

5. Conclusion

There were many studies concluded that global sea level is rising at an increasing rate. While sea level changes are a relatively slow process, therefore, understanding past sea level is important for the analysis of current and future sea level changes. Sea level changes are driven by a variety of mechanisms operating at different spatial and temporal scales [19]. Furthermore, forecasting future relative sea level changes at specific locations requires not just an estimate of global mean sea level changes, but also estimates of the different processes contributing to global mean sea level changes, as well as of the processes contributing exclusively to regional or local mean sea level changes [20].

Forecast is a kind of dynamic filtering, in which past values of the time series are used to predict future values. This study not only wanted to find the best fit model for the time series, but also tried to evaluate the accuracy of the SARIMA, NARNN-LM, and Mixed-LM models used in the forecast of monthly mean sea level at Grand Isle, Louisiana. The comparative results revealed that the Mixed-LM model with 9 neurons in the hidden layer and 3 time delays (MSE = 2.29385e-3) yielded higher accuracy than the NARNN-LM model with 9 neurons in the hidden layer and 36 time delays (MSE = 2.32078e-3), and the ARIMA(1,1,1)$(2,0,0)_{12}$ with drift model (MSE = 2.45154e-3) in this study.

Empirically, the SARIMA and NARNN models are good at modelling linear and nonlinear problems for the time series, respectively. However, using the Mixed model, a combination of the SARIMA and NARNN models has both linear and nonlinear modelling capabilities, can be a better choice for modelling the time series. According to the results of this study, this Mixed SARIMA-NARNN model not only can provided richer information which are important in decision making process related to the future local mean sea level impacts, but also can be employed in forecasting the future performance for local mean sea level change outcomes.

References

[1] M. H. Beale, M. T. Hagan, and H. B. Demuth, Howard B, "Deep Learning Toolbox™: Getting Started Guide," Natick, MA: The MathWorks, Inc., 2019.

[2] G. Benrhmach, K. Namir, A. Namir, and J. Bouyaghroumni, "Nonlinear autoregressive neural network and extended Kalman filters do prediction of financial time series," Journal of Applied Mathematics, Vol. 2020, Article ID 5057801, 1-6, 2020.

[3] D. Bolin, P. Guttorp, A. Januzzi1, D. Jones1, M. Novak1, H. Podschwit, L. Richardson, A. S¨arkk¨a, C. Sowder, and A. Zimmerman, "Statistical prediction of global sea level from global temperature," Statistica Sinica, 25, 351-367, 2015.

[4] G. E. O. Box, and G. M. Jenkins, "Time series analysis: forecasting and control," Holden-Day, San Francisco, 1970.

[5] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, "Time series analysis: forecasting and control (5th ed.)," Hoboken, N.J.: John Wiley and Sons Inc., 2016.

[6] A. Braakmann-Folgmann, R. Roscher, S. Wenzel, B. Uebbing, and J. Kusche, "Sea level anomaly prediction using recurrent neural networks," In Proceedings of the 2017 Conference on Big Data from Space, pp. 297-300, 2017.

[7] N. Bruneau, J. Polton, J. Williams, and J. Holt, "Estimation of global coastal sea level extremes using neural Networks," Environmental Research Letters, 15(7), 074030, 1-11, 2020.

[8] A. Cazenave, and W. Llovel, "Contemporary sea level rise," Annual Review of Marine Science, 2, 145-173, 2010.

[9] A. Cazenave, and G. L. Cozannet, G.L. "Sea level rise and its coastal impacts," Earth's Future, 2, 15–34, 2013.

[10] J. A. Church, N. J. White, R. Coleman, K. Lambeck, and J. X. Mitrovica, "Estimates of the regional distribution of sea level rise over the 1950-2000 period," Journal of Climate, 17(13), 2609-2625, 2004.

[11] J. A. Church, and N. J. White, "A 20th century acceleration in global sea-level rise," Geophysical Research Letters, 33, L01602, 1-4, 2006.

[12] J. A. Church, N. J. White, T. Aarup, W. S. Wilson, P. L. Woodworth, C. M. Domingues, J. R. Hunter, and K. Lambeck, "Understanding global sea levels: past, present and future," Sustainability Science, 3, 9-22, 2008.

[13] J. A. Church, and N. J. White, "Sea-level rise from the late 19th to the early 21st century," Surveys in Geophysics, 32, 585-602, 2011.

[14] G. Foster, and P. T. Brown, "Time and tide: analysis of sea level time series," Climate Dynamics, 45, 1-2, 291-308, 2014.

[15] H. P. Gavin, "The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems," Department of Civil and Environmental Engineering Duke University, 19 pp., 2020. Retrieved from: http://people.duke.edu/~hpgavin/ce281/lm.pdf

[16] M. Haasnoot, J. Kwadijk, J. Alphen, D. Bars, B. Hurk, F. Diermanse, A. Spek, G. O. Essink, J. Delsman, and M. Mens, "Adaptation to uncertain sea-level rise; how uncertainty in Antarctic mass-loss impacts the coastal adaptation strategy of the Netherlands," Environmental Research Letters, 15, 034007, 1-15, 2020.

[17] B. P. Horton, R. E. Kopp, A. J. Garner, C. C. Hay, N. S. Khan, K. Roy, and T. A. Shaw, "Mapping sea-level change in time, space, and probability," Annual Review of Environment and Resources, 43, 481-521, 2018.

[18] IPCC, "Climate Change 2014: Synthesis Report," Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, PachauriR.K., & Meyer, L.A. (eds.)], IPCC, Geneva, Switzerland, 151 pp., 2014.

[19] R. E. Kopp, B. P. Horton, A. C. Kemp, and C. Tebaldi, "Past and future sea level rise along the coast of North Carolina, USA," Climatic Change, 132, 693–707, 2015.

[20] R. E. Kopp, C. C. Hay, C. M. Little, and J. X. Mitrovica, "Geographic variability of sea-level change," Current Climate Change Reports, 1, 192–204, 2015.

[21] S. A. Kulp, and B. H. Strauss, "New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding," Nature Communications, 10, 4844, 1-12, 2019.

[22] G. M. Ljung, and G. E. O. Box, "On a measure of lack of fit in time series models," Biometrika, 65(2), 297-303, 1978.

[23] O. Makarynskyy, D. Makarynska, M. Kuhn, and W. E. Featherstone, "Predicting sea level variations with artificial neural networks at Hillarys Boat Harbour, Western Australia. Estuarine," Coastal and Shelf Science, 61(2), 351–360, 2004.

[24] D. C. Montgomery, C. L. Jennings, and M. Kulahci, "Introduction to time series analysis and forecasting," Hoboken, N.J.: John Wiley & Sons. Inc., 2008.

[25] P. K. Srivastava, T. Islam, S. K. Singh, G. P. Petropoulos, M. Gupta, and Q. Dai, "Forecasting Arabian sea level rise using exponential smoothing state space models and ARIMA from TOPEX and Jason satellite radar altimeter data," Meteorological Applications, 23, 633-639, 2016.

[27] U.S. Global Change Research Program (USGCRP), "Climate science special report: Fourth National Climate Assessment, Volume I," [Wuebbles, D. J., Fahey, D. W., Hibbard, K. A., Dokken, D. J., Stewart, B. C., and Maycock, T. K. (eds.)], U.S. Global Change Research Program, Washington, DC, USA, 470 pp., 2017.

[28] H. Visser, S. Dangendorf, and A. C. Petersen, "A review of trend models applied to sea level data with reference to the "acceleration-deceleration debate," Journal of Geophysical Research: Oceans, 120(6), 3873-3895, 2015.

[29] W. Wang, and H. Yuan, "A tidal level prediction approach based on BP neural network and Cubic B-Spline Curve with Knot Insertion Algorithm," Mathematical Problems in Engineering, 2018, Article ID 9835079, 9 pp., 2018.

[30] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," Neurocomputing, 50, 159-175, 2003.