Short-term sea level forecasting using machine learning techniques: A case study for South Africa

Presented by:
Maria Kristina Ingreso

Prepared for:
Assoc. Prof Albertus Smit
Dr. Christo Rautenbach
Assoc. Prof Marcello Vichi

Submitted to the Department of Oceanography at the University of Cape Town in partial fulfilment of the academic requirements for a Masters of Science (MSc.)

January 2022

Declaration

I, **Maria Kristina Ingreso**, hereby declare that the work on which this dissertation thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature:... Signed by candidate .....

M. K. Ingreso

Date: January, 2022

Acknowledgments

I would like to thank my supervisors Associate Professor Albertus Smit, Dr. Christo Rautenbach and Associate Professor Marcello Vichi for providing me with the guidance and support throughout my masters.

I would like to express my gratitude to the National Research Fund (NRF) who provided financial assistance for my research.

I am extremely grateful to my parents, Soledad and Romeo Ingreso, and my sisters, Maria Katrina Sobremonte and Maria Korina Ingreso for always being there to provide motivation and support.

Lastly, thank you to my boyfriend, Adam Gild, for all your love and support.

Abstract

Seawater levels along the South African coastline are investigated with the use of machine learning techniques. In this study, data-driven methods, which are more computationally efficient in comparison to numerical models, are applied to predict seawater levels. The open-loop NARX model was developed using the Neural Net Time Series application from the Deep Learning Toolbox 14.0 provided by MATLAB® (Mathworks, 2020). A total of five inputs (atmospheric pressure, mean wave period and direction, wind speed and direction) and a single output of seawater level was fed into the neural network where 70 % of the data was used for training, 15 % was used for validation and the remaining 15 % was used to test the model. Three separate storm events that occurred along the coast of South Africa were used for the final model validation. Model performance was measured using the correlation coefficient (R), the root mean square error (RMSE), the bias and the Willmott indices of correlation. It was found that, through principal component analysis (PCA), atmospheric pressure, wind speed and direction and mean wave period and direction are important physical drivers of sea level. The overall model performance was better when all five met-ocean variables were included as inputs to the model than when one or two were excluded, with R and RMSE values ranging from 0.85 to 0.99 and 4.344 to 100.5 mm, respectively. The study presented here clearly shows an effective methodology to not only demonstrate the high accuracy the model has on seawater level predictions, but also able to further investigate the importance of what each oceanic and atmospheric variable has on the seawater level. The model performance may be affected by frictional shoaling, coastally trapped waves, bathymetry and the local dynamics contributed by Agulhas Current, which were not taken account for in this study and could be incorporated in the model for future research.

Contents

List of Figures

List of Tables

List of Abbreviations

AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Networks
ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
BP	Backpropagation
C3S	Copernicus Climate Change Service
CO_2	Carbon dioxide
CTW	Coastally Trapped Wave
DEA	Department of Environmental Affairs
ECMWF	European Centre for Medium-Range Weather Forecasts
FDD	Fast Delivery Data
FFNN	Feedforward Neural Network
GCN	GLOSS Core Network
GP	Genetic Programming
IPCC	Intergovernmental Panel on Climate Change
MHW	Marine heatwaves
NARX	Nonlinear Autoregressive Network with Exogenous Inputs

NN	Neural Network
PCA	Principle Component Analysis
QC	Quality Control
R	Correlation Coefficient
RMSE	Root Mean Square Error
RQD	Research Quality Data
SST	Sea Surface Temperature
SAWS	South African Weather Service
UHSLC	University of Hawaii Sea Level Center

Chapter 1

Introduction and Literature Review

The South African coastline spans over 3000 km and is susceptible to changing seawater level which can inhibit daily life, endanger structures, infrastructures, and commercial operations through coastal erosion and inundation (Hetzel et al., 2019; IPCC, 2019; Rautenbach et al., 2020). In the decades ahead, forecasted seawater level rise as a result of climate change is expected to amplify the problems that seawater level rise brings today (IPCC, 2019; Kirezci et al., 2019). Therefore, it is a pressing concern for South Africa to develop a dependable and computationally efficient prediction tools that can aid in forecasting seawater levels as a function of driving metocean variables (surface pressure, mean wave period and direction and wind speed and direction). In the recent years, studies have seen a development and widespread adoption of physics-based numerical models to predict seawater levels.

Such models are powerful and can easily adapt to a number of situations where bathymetric data and boundary condition time series are known, and present answers based on rendering the physics of the underlying processes. However, these models are computationally expensive to run at large spatial scales which is

needed for nation-wide seawater level predictions (Tadesse et al., 2020). Lately, data-driven modelling approaches have gained high recognition because in contrast to numerical models, data-driven models are more computationally efficient and cheap and does not require large amounts of data to run (Cid et al., 2018; Shamshirband et al., 2020). This thesis is about developing data-driven models to predict seawater levels, based on machine learning techniques, along the coast of South Africa which can be used to aid current South African forecasting systems.

The thesis is structured as follows; Chapter 1 provides a detailed overview of climate change, local and global extreme events, and a comprehensive study on artificial intelligence and relevant machine learning techniques. Chapter 2 describes the data and methods. Chapter 3 presents the results that were obtained from the machine learning techniques and will discuss the results from the tested models and will review relevant patterns to address the aims and objectives and lastly, Chapter 4 will summarise the final conclusions of the study.

1.1 Climate change and extreme events

The concentration of atmospheric carbon dioxide (CO_2) has been increasing and is continuing to do so (Lemonnier et al., 2018) over the last century due to the burning of fossil fuels. Before the pre-industrial revolution, the atmospheric CO_2 concentration was approximately 280 ppm (Le Quéré et al., 2018; Lemonnier et al., 2018). Today, the global concentration value of CO_2 has exceeded 414 ppm (Figure 1.1) (Dlugokencky and Tans, 2018; Fig. 1, Le Quéré et al., 2018; Lemonnier et al., 2018; Lindsey, 2020, IPCC, 2021). The additional anthropogenic CO_2 to the atmosphere is the main driver of climate change because of the greenhouse effect (Meehl et al., 1999; Stocker et al., 2013). A key concern with climate change is the increase in frequency, magnitude, and intensity of extreme events (Mason et al., 1999; Mirza, 2003; Beniston and Stephenson, 2004; Wernberg et al., 2012; Trenberth et al., 2015; Bellprat and Doblas-Reyes,

2016; Hobday et al., 2016). There is evidence that suggests that some extreme events have increased globally since 1950, specifically heavy precipitation, the frequency of high temperatures, and high seawater levels (Herring et al., 2014; Chen and Sun, 2015; IPCC, 2014; Stott et al., 2016, IPCC, 2021).

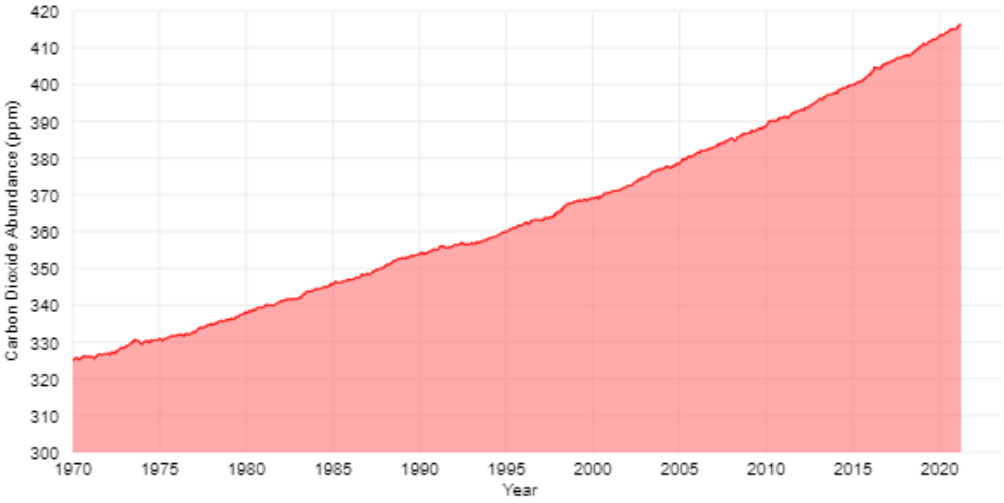


Figure 1.1: The average concentration of CO_2 recorded at NOAA's Mauna Loa Observatory situated in Hawaii (Lindsey, 2020).

Extreme events (e.g. heat waves, droughts, hurricanes, storm surge, and tropical cyclones) can be described based on the following: frequency, their intensity, and the impacts which they have on the environment and/or economy (Easterling et al., 1999; Meehl et al., 1999; Rosenzweig et al., 2001; Mirza, 2003; Beniston and Stephenson, 2004; Diez et al., 2012; Hobday et al., 2016; Nizamani et al., 2017; IPCC,2021). In 1988, the World Meteorological Organization and the United Nations Environment Programme set up the Intergovernmental Panel on Climate Change (IPCC). This was done to supply policy makers with assessments regarding climate change on a regular basis (i.e. what the impacts of climate change are, the risks it poses in the future, and ways to adapt and mitigate climate change). The Fifth Assessment Report by the IPCC defined an extreme weather event that is rare at a certain area and time of the year. Essentially, rare in this case means that the event has to be as rare as or rarer than the

upper and lower 10 % of a probability density function which is estimated from the observations (IPCC, 2014). The increasing amount of compound extreme events (which is a combination of hazards and/or various drivers that contribute to the environmental or societal risk) is likely due to human influences (IPCC, 2021).

1.1.1 Precipitation and flooding

Floods are defined as the overflowing of standard limits of a body of water or stream over areas that are not typically submerged which are a result of abnormally heavy rain, storms, and cyclones (IPCC, 2019). The global and regional precipitation is expected to intensify (Donat et al., 2016) and precipitation extremes are of concern because of their potential economic and social impacts (O'Gorman, 2015). The warming climate leads to enhanced evaporation from the surface, resulting in intense periods of drought (Trenberth, 2011). According to the Clausius-Clapeyron relationship, precipitation is expected to intensify as the atmosphere's ability to hold more water increases by about 7 % per °C warming (Prein et al., 2017; Trenberth, 2011; Wang et al., 2017). Precipitation extremes that occur for short periods can result in flash floods, landslides, and run-off which promote drier soil (Trenberth et al., 2011; Prein et al., 2017). In addition to intensified precipitation, the warming climate would also increase the chance of floods (Hirabayashi et al., 2013). Flooding can also be caused by storm surges which are induced by tropical cyclones or low-pressure systems that are found away from the tropics (Hetzel et al., 2019; Wright et al., 2019), heavy rainfall, or waves (Lee, 2005; Weisberg and Zheng, 2006; Hetzel et al., 2019; Sahoo and Bhaskaran, 2019). Flooding events along the coast often occur coincidentally e.g. a moderate storm surges which occur during a local high tide at the peak of a spring-neap tide cycle (Hetzel et al. 2019) which can have detrimental effects on local coastal communities.

1.1.2 Marine heat waves

Marine heatwaves (MHW) are also predicted to increase in frequency and magnitude (Schlegel et al., 2017). MHWs can be defined as a region that experiences extreme warm temperatures for days to months in the surface waters of the ocean (Hobday et al., 2016). Combinations of atmospheric and oceanographic processes are the cause of MHWs and they have significant impacts on the structure and function of aquatic ecosystems (Oliver et al., 2021).

An atmospheric heatwave during the summer of 2003 in Europe, resulted in amplified air-sea heat flux in the north of the Mediterranean Sea (Schlegel et al., 2017). Weak winds together with the heatwave, induced regional scale thermal stratification and 2 to 3 °C sea surface temperature (SST) anomalies (Garrabou et al., 2009). This led to negative ecological impacts such as the death of benthic invertebrates (Frölicher et al., 2018). In the austral summer of 2011, Western Australia experienced a MHW which was caused by a combination of oceanographic and atmospheric processes related to the strong La Niña of 2010/11 (Wernberg et al., 2021). This La Niña event resulted in the anomalous advection of warm tropical waters towards the poles and into the temperate regions (Oliver et al., 2021). The impacts of this MWH resulted in significant shifts in the ecosystem structure and function of the benthic community in the tropical-temperate transition zone which led to the death of many cold-water species (Wernberg et al., 2012; Smale and Wernberg, 2013; Wernberg et al., 2021). The North-Western parts of the Atlantic Ocean experienced a MHW in 2012, which affected numerous commercial fishing industries (Schlegel et al., 2017). According to the Sixth Assessment Report by the IPCC, most of the MHWs that occurred over 2006 - 2015 have been caused by anthropogenic warming (IPCC, 2021). The frequency of MHWs will continue to increase in the future and as a result, will be an increasing hazard to marine ecosystems (IPCC, 2021).

1.1.3 Extreme winds and waves

Extreme winds create immense amounts of destruction and are a concern for many commercial industries (Outten and Esau, 2013; Kumar et al., 2014) and therefore the understanding of extreme winds in particular areas is important for the design and construction of structures that are exposed to the wind (Outten and Esau, 2013; Kumar et al., 2014). These events can have effects on the energy sectors and coastal ecosystems (Kumar et al., 2014). The increase in greenhouse gas concentrations in the atmosphere causes the continents to warm faster than the oceans and as a result, coastal regions experience an intense pressure gradient between the land and the ocean (Iles et al., 2012). The amplified pressure gradient lengthens the upwelling season which leads to less frequent upwelling events (Iles et al., 2012). Although amplified pressure gradients are assumed to affect upwelling, recent studies (Abrahams et al., 2021) show otherwise, with no changes in the wind patterns resulting in other oceanographic processes affecting upwelling.

Extreme waves relay tremendous amounts of energy onto the coast which exert great stress on coastal structures that provide protection (Mentaschi et al., 2016). These waves are usually formed by strong winds that blow for long periods over long ocean fetches (Hansom et al., 2015), resulting in the acceleration of sediment transport along the coast and coastal erosion (Rodriguez-Ramirez et al., 2015; Archetti et al., 2016; Mentaschi et al., 2016). The changing wave conditions caused by climate change have negative influences on the ecosystem such as coral reefs which protect coastal areas from hazards such as flooding, coastal erosion produced by the dissipation of wave energy, and overtopping (Grady et al., 2013; Ferrario et al., 2014; Quataert et al., 2015). Extreme waves generally arise during storm events that occur at sea, where the waves experience constructive interference resulting in waves that are much larger than the original waves (Hansom et al., 2015). Tsunamis and meteotsunamis (an atmospheric disturbance) can also produce extreme waves (Hansom et al., 2015; Monserrat et al., 2006; Vilibic and Sepic, 2009).

1.1.4 Coastal water level

The interface where land meets the sea is called the coastal zone which is an area that is influenced by atmospheric, marine, and terrestrial processes (Ray, 1991; Hetzel et al., 2019). Coastal environments are important because it serves as natural habitats for plants and animals, and they provide areas that are ideal for commercial operations and human settlement (Ferrario et al., 2014; Hetzel et al., 2019).

The effects of climate change put coastlines at risk globally (i.e. sea level rise) and at a regional scale (i.e. storms and wave climate) (Masselink et al., 2015). Satellite altimeter data which were collected over a 25 year period (starting from 1993) have shown that there is a rise in the global mean sea level ($\sim 3\pm 0.4$ mm/y) (Chambers et al., 2017; Chen et al., 2017; Nerem et al., 2018). The rise in extreme and mean seawater level due to climate change (Kirezci et al., 2020) endangers coastal zones through a number of coastal hazards including: land being permanently submerged by higher mean seawater levels or mean high tides, coastal floods with increasing frequency and intensity, amplified coastal erosion, the loss and transformed coastal ecosystems and the detrimental socio-economic impacts to the coastal communities in low-lying areas (IPCC, 2019). The main processes which lead to global mean sea level rise are the addition of water into the oceans by the melting of ice sheets, ice caps, and glaciers and thermal expansion which causes volume change as the oceans warm, and the alteration of water storage on land (Levitus et al., 2012; Gardner et al., 2013; Chambers et al., 2017; IPCC, 2019). For example, the glaciers that fall outside of Greenland and the Antarctic ice sheets are significant contributors to sea level change due to their ablation and specific accumulation rates, which are notably higher than the ice sheets and are sensitive indicators of climate change (IPCC, 2019).

In coastal areas, extreme seawater level events are defined as the change in water level heights that are brought about by contributing factors which consist of the mean sea level, the bathymetry, storm surge (a temporary rise in the height of

the sea at a particular time and place as a result of extreme meteorological conditions) and tides (Lee, 2005; Merrifield et al., 2014; Hetzel et al., 2019; IPCC, 2019).

Recent literature shows that the rising mean sea level rise is the cause of extreme seawater level events at the coast which bring about coastal floods (IPCC, 2019). Other contributors to rising sea levels are resonant standing waves, river flooding due to heavy rainfall, tsunamis (seismic and meteorological), and coastally-trapped waves which are remotely forced, and other local processes that can result in significant sea level changes at the coast (Hetzel et al., 2019). Therefore, the need for accurate predictions and observations along the coast is crucial for sustainable coastal management (Hetzel et al., 2019).

1.2 Extreme events in South Africa

The Department of Environmental Affairs (DEA) has pointed out the increase of extreme weather events in South Africa (Department of Environmental Affairs, 2018; Climate risk country profile: South Africa, 2021). Figure 1.2 shows the most common types of climatic events that have occurred and been recorded since 1980. These climatic events include floods, storms, wildfires, extreme temperatures, and droughts (Figure 1.2).

In March 2007, a storm swell coincided with a spring tide high on the coast of KwaZulu-Natal. The storm gave rise to 8.5 m swells which resulted in coastal erosion and destruction to the properties along the coast. Coastal modification such as seawalls, the type, shape, and profile of the coast, and poor town planning in the region are some notable local factors that have worsened the effect of coastal erosion (Figure 1.3) (Smith et al., 2007). In 2008, a storm surge event caused by a mid-latitude cyclone and low pressure, occurred on the west and south coast of South Africa which damaged properties and infrastructure near the coast (Davis-Reddy and Vincent, 2017).

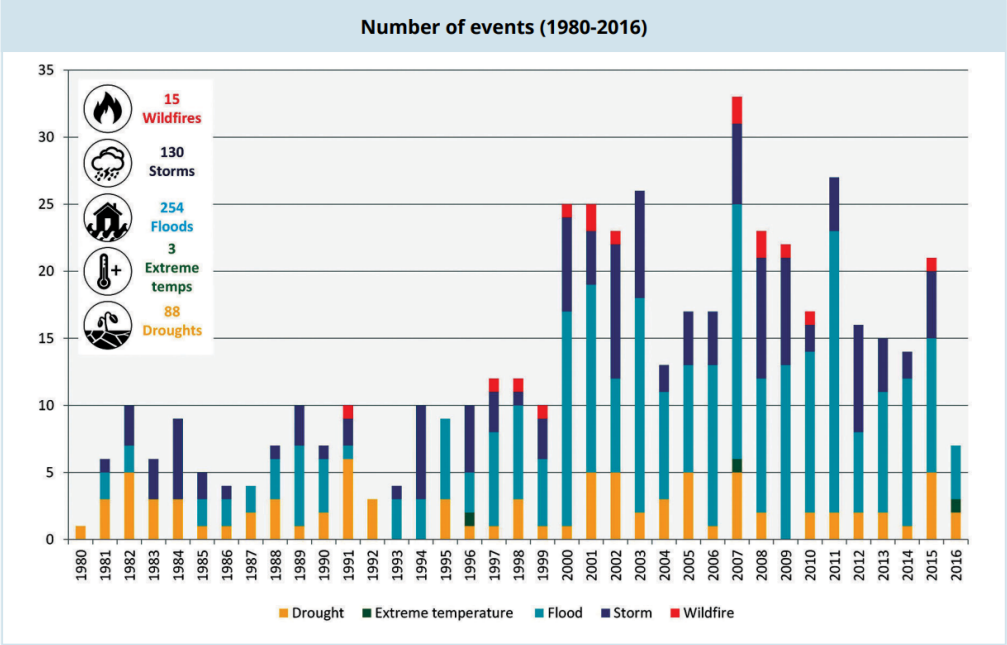


Figure 1.2: The number of recorded events which are related to climatic events (wildfire, floods, storms, extreme temperatures and drought) from 1980 to 2016 (Davis-Reddy and Vincent, 2017).

Figure 1.3: Photo taken at Ballito showing the beach erosion (Smith et al., 2007).

In June 2017, Cape Storm, an extreme storm event that occurred produced violent wind speeds as high as 120 km/h, storm surge and wave heights of 9 to 12 m resulting in thousands being displaced due to floods, eight deaths, and damage to 135 schools.

In July 2020, a cold front brought about gale-force winds blowing at 65-100 km/h

Figure 1.4: Al Jazeera article on Cape Storm in June 2017 (Storm kills several, displaces thousands in Cape Town, 2020).

(Farber et al., 2020) which led to heavy rains, large waves with wave heights of 6–13 m, uprooted trees, floods, and power outages in certain areas (Farber et al., 2020; Mean Geach, 2020). Many areas were affected by this storm, mainly informal settlements situated in Du Noon. Houses were destroyed in Manenberg, Lavender Hill, and Heideveld due to the strong winds (Farber et al., 2020; Geach, 2020).

Figure 1.5: Daily Maverick article on the winter storm surge that occurred in July 2020 in Cape Town (Mean Monday: Another vicious winter storm surges through City of Cape Town, 2020).

1.3 Machine Learning to predict extreme events

1.3.1 Artificial Neural Networks

Artificial Neural Networks (ANN) are inspired by biology (Haupt et al., 2008; Cao et al., 2012) and is one of the many Artificial Intelligence (AI) techniques where the results of its "intelligence" comes from the communication between the different "neurons" (Jain and Deo, 2006). An ANN is composed of a number of small processing units called "neurons" or "nodes", which are individually interconnected (More and Deo, 2003; Jain and Deo, 2006; Mellit, 2008). Generally, an ANN is made up of several layers of interconnected neurons. Each neuron is connected to other neurons in the succeeding layer (Mellit, 2008). Data are introduced to the ANN through an input layer and the output layer holds the network's response to the input. There may be one or more hidden layers that exist between the input and output layer (Mellit, 2008; Xu et al., 2021).

A Neural Network (NN) architecture describes the patterns of connections between the neurons. There are several architectures available e.g. a single layer

NN architecture. The way a NN behaves is defined by how the individual neurons are connected and the strength of the connections or weights between them. The outputs of the network are correlated to the inputs via the neurons with weights and bias. During the training process, the weights are automatically adjusted until it accomplishes the desired error rate (Huang et al., 2003; Xu et al., 2021). A single-neuron model may be easier to understand the NN architecture (Figure 1.6).

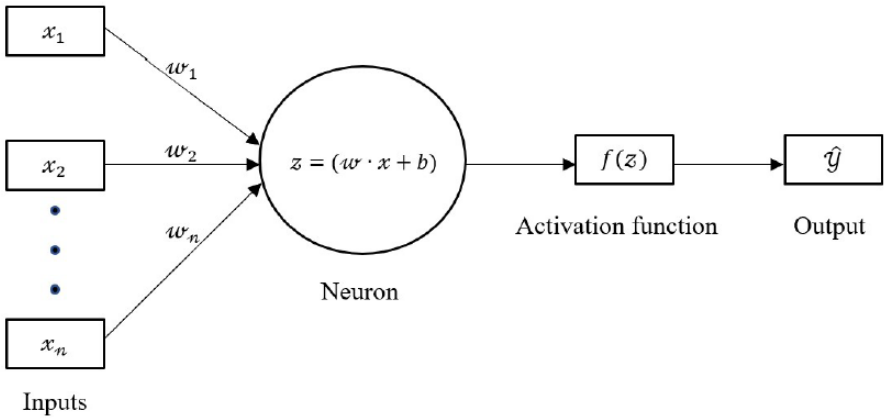


Figure 1.6: Structure of a single neuron NN.

A neuron is an essential unit which processes information throughout the network. Figure 1.6 shows three basic components in an ANN:

(a) *Weights* ($w_i = [w_1, w_2, ..., w_n]$) which are the connecting links between the neurons.

(b) A *hidden layer*, which is the computational layer. Here, the row vectors of the inputs ($x_i = [x_1, x_2, ..., x_n]$ and the weights ($w_i = [w_1, w_2, ..., wn]$) are multiplied and summed.The summation is therefore equal to the dot product of the input and weight vectors.

$$\sum_{i=1}^{n} = (x_i \cdot w_i) \tag{1.1}$$

(c) *Activation functions.* These are functions which introduce non-linearity to the neurons and have an impact on how fast the NN learns. The activations functions associated to the study are:

(i) Linear function:

$$f(z) = z \tag{1.2}$$

(ii) Sigmoid function:

$$f(z) = \frac{1}{1 + e^{-z}} \tag{1.3}$$

(iii) Tanh function:

$$f(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}} \tag{1.4}$$

(iv) Rectified linear unit:

$$f(z) = max(0, z) \tag{1.5}$$

A common type of ANN is a multi-layered perceptron feedforward neural network (FFNN). A FFNN consists of an input layer, one or more hidden layers, and an output layer (Faris et al., 2016). A training algorithm called backpropagation (BP) is a popular gradient-based training algorithm that updates the weights and biases until the performance criteria are satisfied (Montana and Davis, 1989; Faris et al., 2016; Xu et al., 2021). BP succeeds in training simple problems however, as the dimensionality increases or the more complex the data the performance decreases (Montana and Davis, 1989).

1.3.2 Nonlinear Autoregressive Network with Exogenous Inputs

A Nonlinear Autoregressive Network with exogenous inputs (NARX) is a recurrent dynamic network that is applicable for modelling nonlinear dynamic systems (Di Piazza et al., 2016) and was applied in this study to predict seawater levels. In contrast to a FFNN where the flow of information is forward and only in one direction, the information in the NARX network flows in both the forward and backward directions which allows the connection between the neurons to exist

in the same or previous layers (Di Nunno and Granata, 2020). The equation for the NARX model is:

$$y(t) = f(y(t-1), y(t-2), ..., y(t-n_y), u(t-2), ..., u(t-n_u)) \tag{1.6}$$

where the input and output values are *u(t)* and *y(t)* respectively at time *t*, *f* is the non-linear function and the input and output layers are n_u and n_y respectively. The network consists of three layers: the input, hidden and output layers where the atmospheric and oceanic variables are fed into the input layer with tapped delays which incorporate a sigmoid activation function (Equation 1.3) in the hidden layer and a linear activation function in the output layer (Equation 1.2).

1.4 Machine learning in oceanography

This section mainly focuses on machine learning techniques that have been applied to the main ocean and atmospheric drivers of seawater level. Accurate predictions of the inputs to the NN (wind, wave, and surface pressure) need to be made to successfully predict extreme seawater levels. The studies mentioned in the next following subsections for wave, wind, and seawater levels have made use of machine learning techniques which, in general, need much less computational efforts and input variables whilst producing accurate predictions. This results in cost-effective and efficient alternatives to predict the ocean and atmospheric variables of interest. However, models that use machine learning techniques are required to be site-specific and do not take the physical phenomena into account when making predictions. In contrast, numerical models can be applied to any point or region whilst also taking the physical processes into account when making predictions (Shamshirband et al., 2020).

1.4.1 Ocean waves

Information on ocean waves is helpful because of oceanic operations, e.g. repairs on structures or the laying of submarine pipelines, can be planned carefully based on the wave predictions (Agrawal and Deo, 2002; Jain and Deo, 2008; Gopinath and Dwarakish 2015; Lou et al., 2021; Tapoglou et al., 2021).

Deo et al. (2001) developed a FFNN with three layers to forecast waves with lead times ranging from 3 to 24 h. The input of generating wind speeds was used in order to get the output of the significant wave height and the average wave period (Gopinath and Dwarakish, 2015). The method proposed by Deo et.al (2001) is site-specific and is significant when online wave data is being collected at that site. Nevertheless, the results from the NN were compared to the traditional statistical autoregressive models. It was concluded that the NN demonstrated was able to predict wave heights much better than the conventional statistical models with correlation coefficients of 0.52, 0.68, and 0.77 for sites one, two, and three respectively. The low correlation coefficient for site 1 was possibly because of using wind data that was measured at a station that was shore based rather than using wind data that was much closer to the wave rider which was several milometers away (Deo et al., 2001).

Agrawal and Deo (2002) found that NNs performed better than stochastic models (autoregressive integrated moving average (ARIMA) and autoregressive moving average (ARMA)) when the prediction intervals were shorter (3 to 6 hours). The correlation coefficient for the ARMA and ARIMA for 3 h, 6 h and 12 h prediction intervals were 0.803 and 0.821; 0.808 and 0.824; and 0.783 and 0.783 respectively. Whereas the correlation coefficient for the NNs were 0.985, 0.964 and 0.788 for 3 h, 6 h and 12 h prediction intervals.

There are many studies that were done in the past on wave forecasting, most of which make use of NNs. Tsai et al. (2002) reported the use of FFNN to predict the height of the waves at a site. This was based on the observations for wave

data of other sites at Taichung Harbour (Gopinath and Dwarakish, 2015). The wave heights and period forecast was conventionally done using the relationship between the wind and waves however, the error in the coast and harbour area may become large due to the uncertainty in the wind generation prediction in those particular areas (Tsai et al. 2002).

A recent study done by Lou et al. (2021) explored the use of a long short-term memory network (based on recurrent NNs) in order to propose a new type of automatic driving scheme that links ship driving to the prediction of wave heights. Accurate wave height predictions will assist ships in adjusting their course in real-time to allow them to travel in areas with the lowest wave heights. This enhances comfort, safety, and fuel economy. Lou et al. (2021) concluded that the method proposed is applicable to offshore and open sea wave height predictions. The one-hour ahead prediction for wave height yielded a R-value of 0.986 between the predicted and observed wave height values.

Another study done by Tapoglou et al. (2021) trained ANNs to predict significant wave heights. This was done by combining a total of 240 satellite images with wave buoy data to predict significant wave heights to facilitate the safe maintenance of offshore wind farms. It was shown that the method proposed in this study was an effective technique to predict significant wave heights, yielding a RMSE of 0.23 m for wave heights less than 3 m.

1.4.2 Wind

Wind, along with atmospheric pressure and waves, is one of the contributing factors to the seawater level (Muslim et al., 2020; Rautenbach et al., 2020). A study done by Muslim et al. (2020) investigated the use of AI methods to predict the impact of metrological parameters on sea level. Observed values of wind speed and direction were fed as inputs into an adaptive neuro-fuzzy inference system (ANFIS) (a type of ANN). The ANFIS exhibited an accurate performance in training (R and RMSE values ranging from 0.50–0.82 and 0.0070–

0.0086) and testing (R and RMSE values ranging from 0.42–0.80 and 0.0059–0.0095) for the prediction of seawater level and the results from the study has proven to produce a reliable prediction for increases in sea level in coastal areas.

A recent study by Balogun and Adebisi (2021) investigated the use of three different types of trained neural networks to predict sea level which made use of ARIMA, long short-term memory NNs, and support vector regression models where ocean and atmospheric variables were fed into the model (including wind speed) to predict sea level. It was found that the mean R values for the ARIMA, long short-term memory NNs and the support vector regression models were 0.710, 0.748, and 0.853 respectively. It was concluded that the long short-term memory NN was the most effective method in sea level prediction compared to ARIMA and support vector regression models.

1.4.3 Extreme Sea level

It is important to understand and predict seawater levels as environments in the coastal zone are exposed to a variety of marine and land-based hazards such as storm surge, shoreline erosion, flooding, and more (Hamlington et al., 2015; Kopp et al., 2015; Neumann et al., 2015; Muis et al., 2016). Extreme seawater level events are the changes in water level heights that are brought about by contributing factors which consist of the mean sea level, storm surge (a temporary rise in the height of the sea at a particular time and place as a result of extreme meteorological conditions) and tides (IPCC, 2019).

A study done by Makarynskyy et al. (2004), developed and validated an ANN that involved short (hourly seawater level predictions for the next 24 h) and medium-term (mean seawater level predictions for the next half-day, day, 5-day and 10 day periods) sea level forecasting in Western Australia. The results were satisfactory with correlation coefficient ranges from 0.7–0.9 and it was concluded that NN can be applied successfully to other coastal areas given that each NN is site-specific and is suitably trained and validated (Makarynskyy et al., 2004).

Another study by Ghorbani et al. (2010), forecasted sea levels for 12-h, 24-h, 5-day, and 10-day time intervals using observed sea levels in Hillary's Boat Harbor, Australia. The AI techniques used in the study (genetic programming (GP) and ANNs) performed satisfactorily (with correlation coefficients ranging from 0.859–0.558 for ANNs and 0.924–0.873 for GP) and was proved to be an alternative method to forecast sea level other than harmonic analysis which is a conventional approach that is limited because of assumptions that; this methodology to predict seawater levels are data-intensive and years of tidal observations need to be collected and processed to obtain valid seawater level estimates (Ghorbani et al., 2010). Another limitation with harmonic analysis is that, when there is a new location of interest, there may be areas with scarce tidal observations and this can be problematic because the location may share similar quantities of gravitational actions by the Sun and Moon but differ in other components that can be liable for up to 30 % mismatch between harmonic estimates and measurements (Ghorbani et al., 2010).

Daily seawater level forecast studies have also been implemented in Bulgaria (Pashova and Popova, 2011) where numerous forecasting techniques were used such as; Cascade-Feed-Forward, Multiple Linear regression, ANNs, and more. However, it was found that the ANNs are an effective way to recognise the non-linear relationship between input and output sea levels by recognising the pattern in historical datasets for long and short-term forecasts, reaching a correlation coefficient value of 0.822 for the test dataset (Pashova and Popova, 2011). Karimi et al., (2013) also did a study in forecasting sea levels at the Darwin Harbor in Australia. Predictions for sea level were for the next hour, 24-h, 48-h, and 72-h. Three methods were used to forecast, namely; ARMA, ANN, and ANFIS. It was concluded that the ANN and ANFIS methods produced similar R^2 results of 0.964, 0.870, and 0.742; and 0.962, 0.866 and 0.731 for 24-h, 48-h, and 72-h forecasts respectively and performed much better than ARMA which produced R^2 values of 0.925, 0.734 and 0.491 for 24-h, 48-h, and 72-h forecasts.

A study done by Cid et al. (2018) made use of multiple linear regression models

for the time period 1866 to 2012 in Asia to reconstruct daily maximum storm surge levels. It was concluded that the statistical models performed similarly or better than a global numerical model, producing correlation coefficients ranging from 0.83 to 0.93 (Cid et al., 2018). Recently, Tadesse et al., 2020 utilised data-driven models to simulate daily maximum storm surge values globally. It was found that the models performed better when simulating surge values in the extratropical and sub-tropical regions than the tropical regions with an average correlation of 0.79 and 0.45 and a RMSE of 7.5 cm and 5.3 cm, respectively (Tadesse et al., 2020). Tadesse et al. (2020) concluded that data-driven models provide an alternative and complementary method of simulating storm surges which is effective and computationally cheap and to be used in conjunction with numerical models which are computationally expensive. A recent study by Rautenbach et al. (2020) made use of dynamic methods to forecast storm surge. Numerical models have been calibrated and validated and as a result have the ability to forecast storm surge along the coast of South Africa. Ocean and atmospheric parameters (wind, wave, and atmospheric pressure) were used to calibrate the model and it was confirmed that the wind plays a significant role in storm surges in South Africa.

1.5 Objectives of this study

The studies mentioned above made use of machine learning techniques in oceanography to forecast certain atmospheric and oceanic variables which is important for coastal management and the safety of coastal communities. It is evident that the previous studies made use of a variety of models that produced good results based on the RMSE and high correlation coefficient values where some were greater than 0.9. It is also noted that most of these studies forecast a maximum of 72 hours with an exception of Rautenbach et al. (2020) and Makarynskyy et al. (2004). Nonetheless, research on machine learning in oceanography is growing since it is cost-effective, more computationally efficient, and does not require much data to run compared to numerical models. This study provides a more

accurate model using neural networks to predict regional seawater levels forecasted over a longer period (one month). This is necessary because it will help to improve the current forecasting systems in place in South Africa and in turn, can assist Disaster Risk Management with adequate forewarning to implement plans when disasters do occur.

Two general methods to model storm surge include data-driven methods and dynamic methods. Data-driven methods involve the use of machine learning or statistical techniques and efficiently utilise historical data to build a relationship between its predictands and predictors (Harris, 1962; Cid et al., 2017). Dynamic methods utilise shallow water equations in order to explain the physical processes which drive storm surge (Harris, 1962; Cid et al., 2017; Shamshirband et al., 2020).

The objectives of this study are to:

- Develop a model using data-driven methods. This model will be trained and validated using machine learning techniques to predict seawater levels along the coast of South Africa.
- Investigate the relationship between seawater level (predictand) and surface pressure, wind speed and direction, and mean wave period and direction (predictors).
- Supply a data-driven method that is computationally efficient and accurate in predicting storm surges in South Africa since numerical models are computationally expensive and need high-quality bathymetric data to simulate surges (Tadesse et al., 2020).

Chapter 2

Data and methods

2.1 Data sources

2.1.1 ERA-5

The European Centre for Medium-Range Weather Forecasts (ECMWF) has released its latest climate reanalysis dataset, ERA-5, which is being developed through the Copernicus Climate Change Service (C3S). The Climate Data Store datasets are evaluated by the Evaluation and Quality Control function of the C3S. This is done independently from the data supplier (Hersbach et al., 2021). The dataset has replaced ERA-Interim as of the 31st of August 2019. ERA-5 provides hourly data on various land-surface, sea-state and atmospheric variables, on 37 pressure levels, over the period 1979 to date. Reanalysis datasets are accessible in the C3S on regular latitude-longitude grids at a horizontal resolution of 0.25° x 0.25° and 0.5° x 0.5° for mean, variance and member datasets. The file formats are available in netCDF and GRIB. Table 2.1 shows the atmospheric and oceanic variables that were used for this study.

Table 2.1: Data downloaded from ERA-5.

Variable	Unit	Horizontal resolution (°)
Surface pressure	Pa	0.25×0.25
10 metre wind speed	m/s	0.5×0.5
10 metre wind direction	degrees	0.5×0.5
Mean wave period	s	0.5×0.5
Mean wave direction	degrees	0.5×0.5

2.1.2 University of Hawaii Sea Level Center

The University of Hawaii Sea Level Center (UHSLC) is one of the major networks of tide gauges that provide tide gauge data globally. Two types of datasets that differ in the level of quality control (QC): Fast Delivery Data (FDD) and Research Quality Data (RQD). FDD receives basic QC which mainly focuses on the large level shifts and apparent outliers. The FDD are issued within 1–2 months of data collection. RQD are regarded to be datasets that are final and science-ready. RQD undergoes a time-consuming RQ process which results in the data being available only 1–2 years after it has been received from the originators by the UHSLC. FDD will be added to the RQD only after the final QC. Many other international agencies contribute to FDD, however, RQD are only available from the UHSLC at daily and hourly temporal resolutions. Although there are numerous tide gauge records in the UHSLC database, the GLOSS Core Network (GCN) serves as the backbone because it has an estimated 300 tide gauge stations globally which is the foundation of the seawater level network (Figure 2.1) (Caldwell et al., 2015).

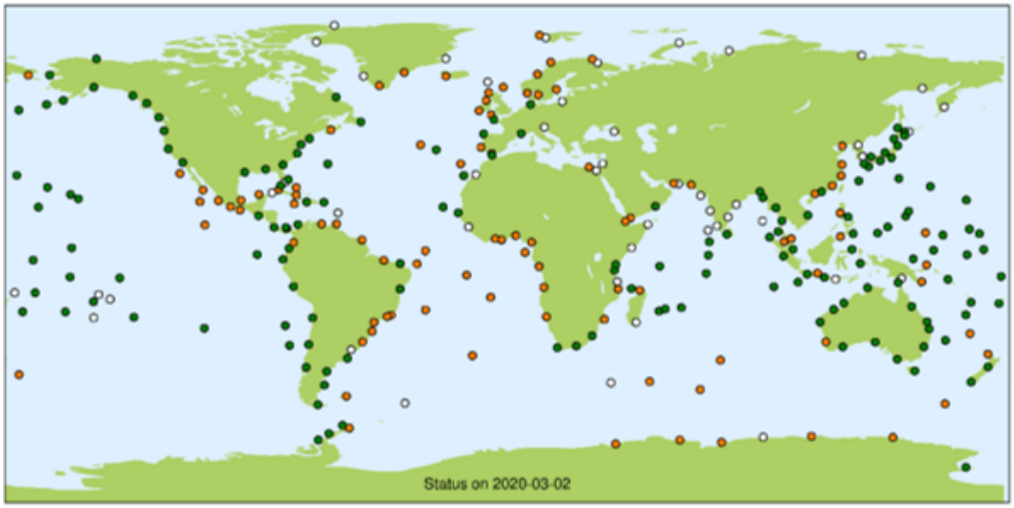


Figure 2.1: Map showing the GCN tide gauge stations where green indicates an operational station, orange indicates that it has been operational in the past and white means it has not been successful in operation (Psmsl.org. 2020. GLOSS Network Status. [online] Available at: https://www.psmsl.org/products/gloss/status.php [Accessed 11 March 2020]).

2.2 Study region and tide gauge locations

There are eight tide gauge stations situated in South Africa (Figure 2.2). A table of the longitude and latitude locations of the tide gauge stations is provided in Table 2.2. Tide gauge data are available on the `http://uhslc.soest.hawaii.edu/thredds/uhslc_quality_hourly.html` website, where both hourly and daily data are available for all eight stations. In this study, hourly seawater level data from UHSLC website was extracted using the legacy data portal (`http://uhslc.soest.hawaii.edu/data/`). The data from the legacy portal were preferred over the data available on the `http://uhslc.soest.hawaii.edu/thredds/uhslc_quality_hourly.html` website since the seawater level data are more up to date, the data can be downloaded as a CSV, NetCDF (newer and older version) and are more user-friendly. A map of South Africa with all the tide gauge locations is provided in (Figure 2.2), however, in this study, one tide gauge from each coast will be used i.e. Simon's Town (west coast), Port Elizabeth (south coast), and Durban (east coast). Another reason for this is that the legacy portal only provides seawater level data for these stations which coincide with Figure 2.1 that show these stations are the only three operational sites.

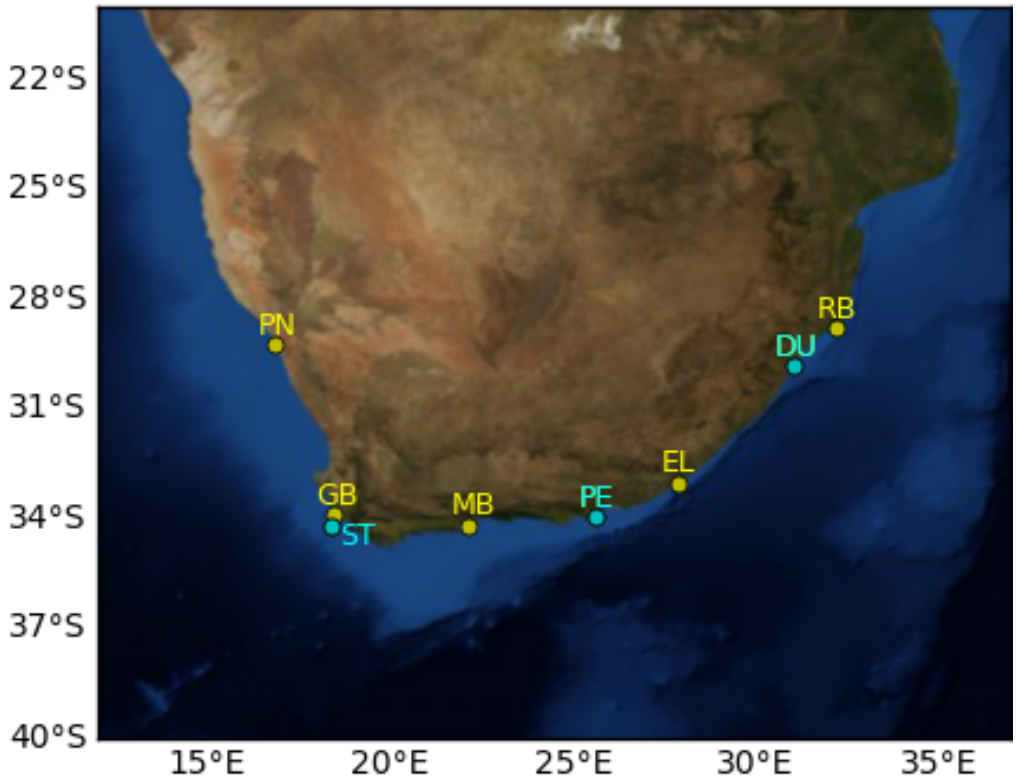


Figure 2.2: Tide gauge stations that are present in South Africa where PN, GB, ST, MB, PE, EL, DU and RB stand for Port Nolloth, Granger Bay, Simon's Town, Mossel Bay, Port Elizabeth, East London, Durban, and Richard's Bay respectively. In this study ST, PE and DU (cyan coloured dots) will be used to validate the model. Although the other stations are not currently operational, seawater level data may still be found at `http://uhslc.soest.hawaii.edu/thredds/uhslc_quality_hourly.html`.

Table 2.2: The longitude and latitude locations of the tide gauge stations in South Africa.

Location	Latitude	Longitude
Port Nolloth	-29.25	16.87
Granger Bay	-34.18	18.43
Simon's Town	-33.90	18.45
Mossel Bay	-34.18	18.43
Port Elizabeth	-33.97	25.63
East London	-33.00	27.90
Durban	-29.83	31.03
Richard's Bay	-28.79	32.18

2.3 Data preprocessing

Data preprocessing was done in MATLAB® (Mathworks, 2020), an interactive software used for numerical computations (Higham and Higham, 2016). The netCDF files from ERA-5 and MS Excel files of hourly seawater level data obtained from UHSLC were imported into MATLAB. Both datasets were combined into one structured timetable where the columns and rows correspond to the variables and observations respectively. The presence of missing values (NaNs) varied from site to site and was mainly due to the mismatch in the horizontal resolution between each of the ocean and atmospheric variables and the percentage of NaNs ranged from 4.12 to 16 %. The percentage of NaNs for seawater level ranged from 0.82–9.8 %. Interpolation was done in MATLAB® (Mathworks, 2020) using the built-in function, retime. A 33-hour low-pass filter (pl66tn) (Beardsley and Rosenfeld, 1983) was applied to the hourly seawater levels to ensure that the tidal signals were removed. Prior to the data being fed into the network, a moving mean was applied to the data using the built-in function in MATLAB® (movmean) to reduce the noise since studies have shown that the presence of noise in a data set resulted in a decrease in the accuracy of model predictions and yielded poor prediction results (Ahmed et al., 2019; Gupta and Gupta, 2019).

2.3.1 Principal component analysis

Principal component analysis (PCA) is an effective multivariate statistical technique that extracts significant information about any latent common structure that a dataset may have. It is very useful, especially when there are large datasets that need to be analysed (Abdi and Williams, 2010; Ait-Sahalia and Xiu, 2015; Xu et al., 2018). PCA was done on sea level, atmospheric and oceanic variables to determine the drivers of seawater level and the strength of the influence that the different variables have on seawater level. Two PCA axes were retained (components 1 and 2), which are plotted on the x and y-axis respec-

tively (Figure 3.1). Despite only using three tide gauge stations to develop, train and validate the model, the PCA will be carried out on all eight sites to confirm the inverse relationship between seawater level and surface pressure and further investigate how wind speed and direction, and mean wave period and direction affect seawater level on the West and East coasts of South Africa. Furthermore, data were available for all eight sites to perform the PCA.

2.4 The neural network model

The NARX model was applied to three different storms that occurred along the coast of South Africa: the storm that hit Durban in March 2007; The Cape Storm of June 2017 and the storm that occurred in July 2020 in Cape Town (Figures 1.3– 1.5). Separate site-specific NARX models were trained for each storm and the number of models per storm is dependent on the availability of data. The table below (Table 2.3) shows the sites used per storm and the time period used to train, test, and validate the storm. The month in which the storm occurred was extracted out of the time period used to train, test, and validate the model using the Neural Net Time Series application from the Deep Learning Toolbox 14.0 provided by MATLAB® (Mathworks, 2020). This serves as an independent data set to see if the model can be utlised for real-world application.

Table 2.3: The number of models that were trained, tested and validated for each storm event. There is only one model per site.

Event	Site	Time Period
June 2017	Durban	2011 - 2017
	Port Elizabeth	2011 - 2017
	Simon's Town	2011 - 2017
March 2007	Durban	2001 - 2007
July 2020	Simon's Town	2014 - 2020

The open-loop NARX model was developed using the Neural Net Time Series application from the Deep Learning Toolbox 14.0 provided by MATLAB® (Mathworks, 2020). The NARX network was trained using supervised learning that utilises a backpropagation training algorithm to adjust the weights between

the neurons in order to minimise the error (Zubier and Eyouni, 2020). A total of five inputs (atmospheric pressure, mean wave period and direction, wind speed and direction) and a single output of seawater level was fed into the neural network (Figure 2.3) where 70 % of the data is used for training, 15 % is used to validate and the remaining 15 % is used to test the model. In MATLAB, the default ratios to split the data into training, testing and validation are 0.7, 0.15 and 0.15 respectively (Mathworks, 2020). The default data division was retained when developing the NARX network since previous studies done by Filippo et al. (2012) has also separated the data into 70 % ,15 % and 15 % when using artificial neural networks to improve sea level forecasting. Nitsure et al. (2014) and Di Nunno and Granata (2020) have also divided the dataset into 70 % training and 30 % testing when predicting sea water levels and groundwater level respectively.

The validation set measures the network's ability to generalise and once the generalisation does not improve, the training ends. The test set provides an independent measure of the network's performance and has no effect on the training. There are three training algorithms available in the Deep Learning Toolbox 14.0 provided by MATLAB® (Levenberg-Marquardt, Bayesian regularization, and the Scaled Conjugate Gradient). Each training algorithm was assessed for each site to determine which one was well-suited for this study in terms of model accuracy and computational efficiency. Table 2.4 shows each station, the correlation coefficient (R), and the time taken to train, test, and validate the model. In this thesis, the Levenberg-Marquardt training algorithm was implemented since it produced accurate model results and it took less time to train, test, and validate the models.

Table 2.4: Comparison between the three training algorithms; Levenberg-Marquardt (LM), Bayesian regularization (BR) and the Scaled Conjugate Gradient (SCG). The correlation coefficient and the time taken for each model to be train, tested and validated is shown.

Storm event	**Site**	**Training algorithm**					
		LM		**BR**		**SCG**	
		R	**Time (s)**	***R***	**Time (s)**	***R***	**Time (s)**
June 2017	Durban	0.99	5	0.99	94	0.98	158
	Port Elizabeth	0.99	82	0.99	1162	0.99	1100
	Simon's Town	0.99	5	0.98	1523	0.98	66
March 2007	Durban	0.99	71	0.99	1200	0.99	350
July 2020	Simon's Town	0.99	5	0.99	35	0.94	82

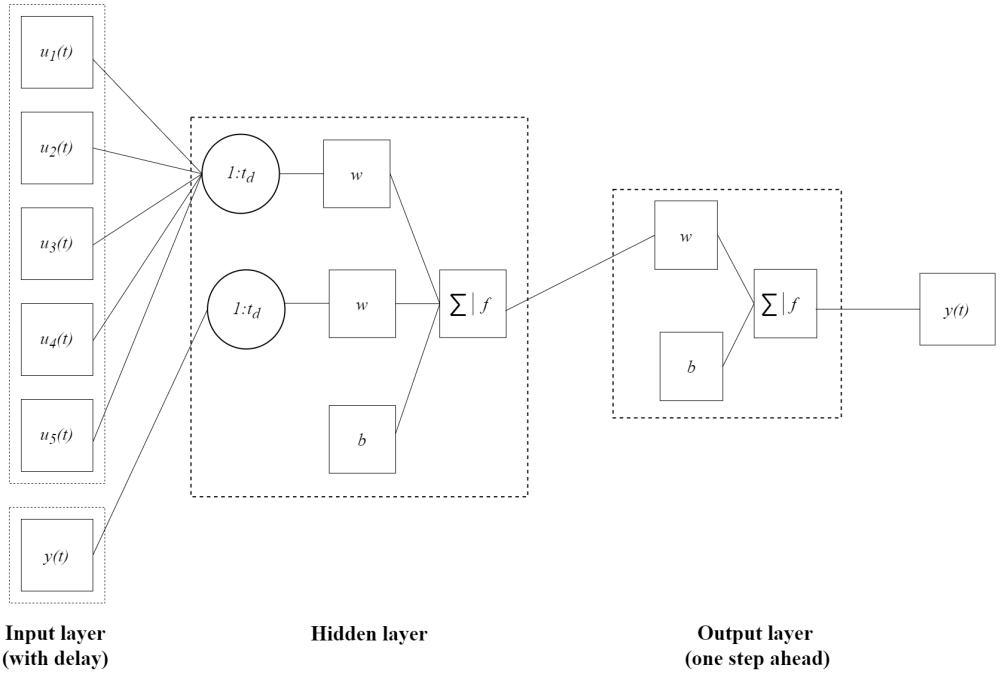


Figure 2.3: A schematic showing the architecture of the NARX network.

2.4.1 Evaluation metrics

After training and validation, the test set was extracted from MATLAB and the performance was measured using the correlation coefficients (R) (Equation 2.1) and the root mean square error (RMSE) (Equation 2.2). The Willmott indices of correlation (Equation 2.3) (Willmott et al., 2012) and the bias (Equation 2.4) were also included as part of the evaluation metrics.

$$R = \frac{\sum_{i=1}^{i=n}(Y_i - \bar{Y})(\hat{Y}_i - \bar{\hat{Y}})}{\sqrt{\sum_{i=1}^{i=n}(Y_i - \bar{Y})^2(\hat{Y}_i - \bar{\hat{Y}})^2}} \tag{2.1}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{n}^{i=1}(Y_i - \hat{Y}_i)^2} \tag{2.2}$$

$$Willmott\ Index = 1 - \frac{\sum_{n}^{i=1}|\hat{Y}_i - Y_i|}{2\sum_{n}^{i=1}|Y_i - \bar{Y}_i|} \tag{2.3}$$

$$Bias = Y_i - \hat{Y}_i \tag{2.4}$$

where: Y_i = Observed seawater level

$\hat{Y}_i$ = Predicted seawater level

$\bar{Y}_i$ = Average of the observed seawater level

$\bar{\hat{Y}}_i$ = Average of the predicted seawater level

Chapter 3

Results and Discussion

Seawater level predictions were made along the South African coastline using machine learning techniques. The NARX models that were developed were applied to the validation set (storm events) which is completely independent from the data set used to train the models. The results presented in the validation set do not include and post-process corrections and were solely based on the model settings during the training phase. Each model is site-specific and thus the model performance will vary from site to site.

3.1 Principal component analysis

PCA results are shown in Figure 3.1 for all eight tide gauge stations. It is evident that the inverse relationship between seawater level and surface pressure exists and furthermore, on the east coast (Port Elizabeth, East London, Durban and Richard's Bay) it shows that wind speed has a weak signal of correlating with seawater level, this can be seen by the large, diverging angle (almost 180°) between the two variables in the PCA biplot between seawater level and surface pressure and the small angle between wind speed and seawater level, respectively. Table 3.1 shows the percentage of variance explained by each Principal Component

for each station. The values for the stations on the west coast and south coast (Port Nolloth, Granger Bay and Mossel Bay) have higher percentage of variance that is explained by Component 1, ranging from 41.22 to 42.46. The stations situated on the south and east coast (Port Elizabeth, East London, Durban and Richard's Bay) have a lower percentage of variance explained by Component 1 compared to the west coast, with values around 30. The variance explained by Component 2 are all similar for all eight stations with values ranging from 19.83 to 23.61.

Table 3.1: Percentage of the variance explained by the fist principal component (PC1) and the second principal component (PC2).

Parameter	**% of variance explained by PC1**	**% of variance explained by PC2**
Port Nolloth	41.22	20.36
Granger Bay	42.07	20.22
Simon's Town	32.75	22.60
Mossel Bay	42.46	19.83
Port Elizabeth	30.98	23.42
East London	30.94	23.17
Durban	30.74	23.61
Richard's Bay	30.74	23.61

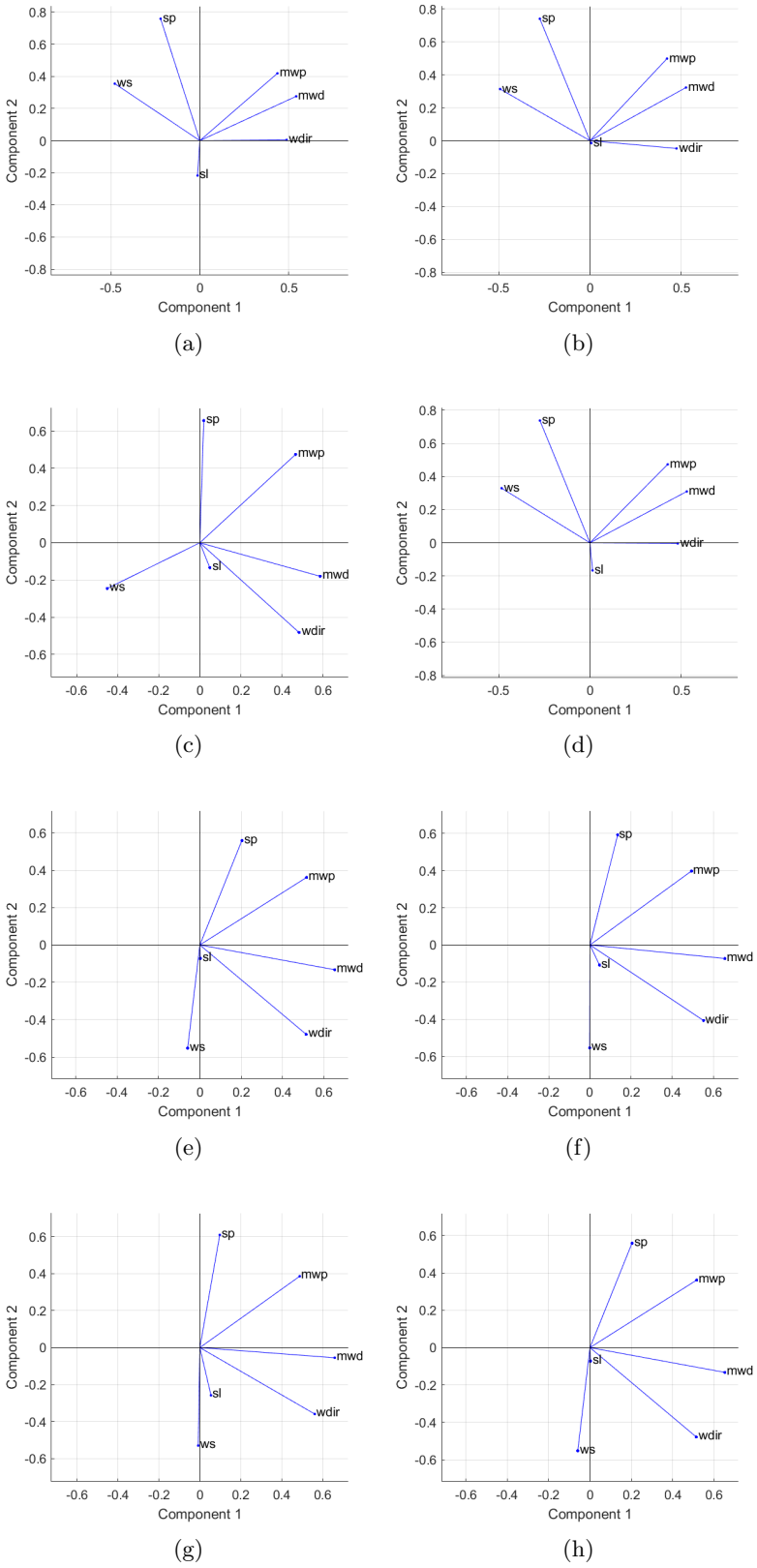
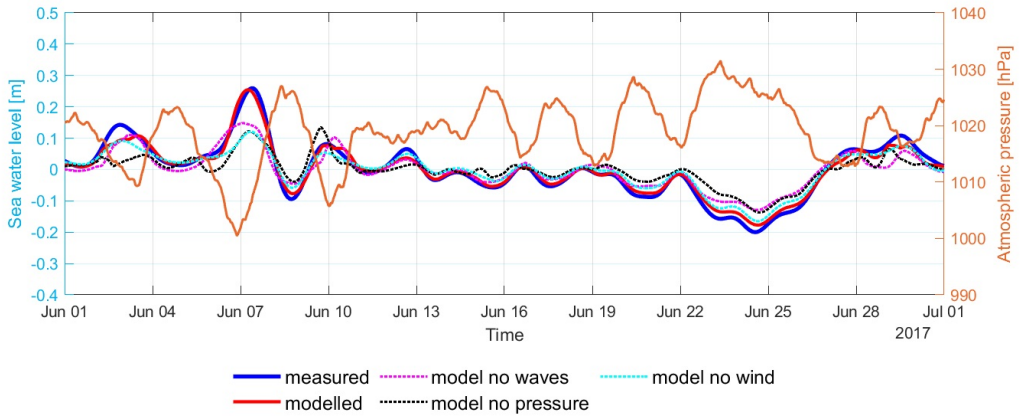


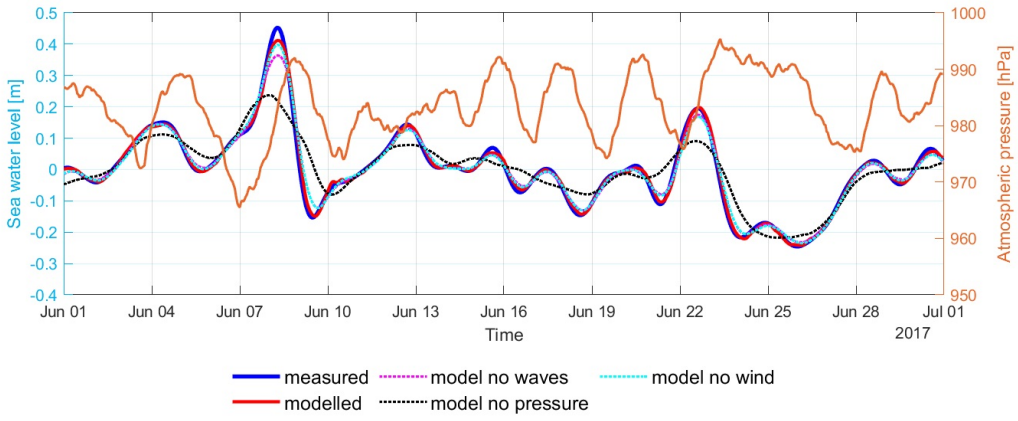
Figure 3.1: PCA done on the measured sea level data together with the influential met-ocean variables (sp is the surface pressure, ws is the wind speed, wdir is the wind direction, mwp is the mean wave period, mwd is the mean wave direction and sl is the sea level) for a) Port Nolloth, b) Granger Bay, c) Simon's Town, d) Mossel Bay, e) Port Elizabeth, f) East London, g) Durban, and h) Richard's Bay.

3.2 NARX model performance

The comparison between the modelled and the measured seawater level, shown in Figures 3.2– 3.4, distinctly indicate the high performance accuracy of the model. The performance was also statistically evaluated using the R and the RMSE values (Table 3.2). The high R and low RMSE values which range from 0.85–0.99 and 4.344–100.5 mm respectively, suggest a strong relationship between the measured and the modelled seawater levels.

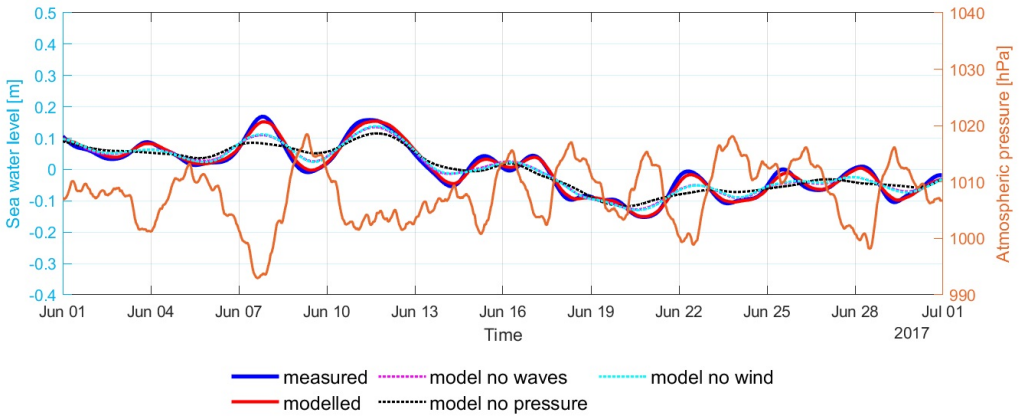


(a)



(b)

Figure 3.2: Final model validation for Cape Storm of June 2017 (a) Simon's Town, (b) Port Elizabeth and (c) Durban. Each panel indicates the measured (blue) and the modelled (red) water levels after the tidal signals have been filtered out on the left axis. The modelled seawater levels without the influence of mean wave period and direction (pink), without the influence of pressure (black) and without the influence of wind speed and direction (cyan) are presented as dashed lines. The atmospheric pressure is provided on the right axis.



(c)

Figure 3.2: Final model validation for Cape Storm of June 2017 (a) Simon's Town, (b) Port Elizabeth and (c) Durban. Each panel indicates the measured (blue) and the modelled (red) water levels after the tidal signals have been filtered out on the left axis. The modelled seawater levels without the influence of mean wave period and direction (pink), without the influence of pressure (black) and without the influence of wind speed and direction (cyan) are presented as dashed lines. The atmospheric pressure is provided on the right axis.

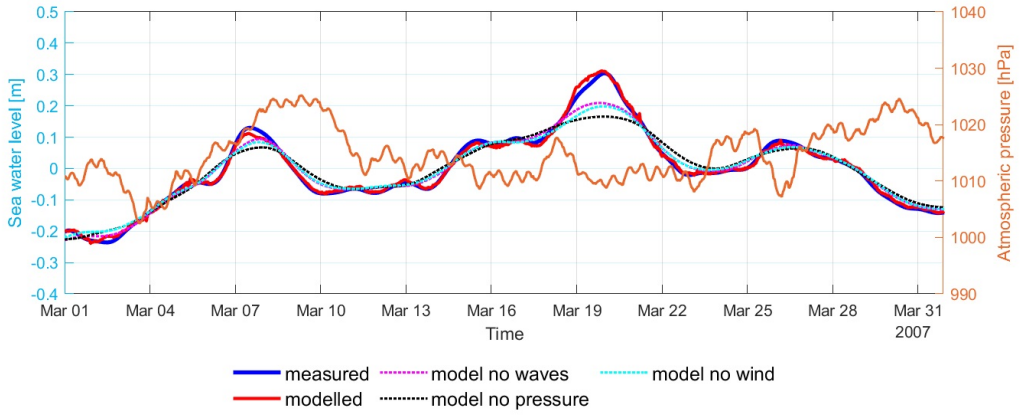


Figure 3.3: Final model validation for the storm that occurred in Durban in March 2007. The blue line indicates the measured seawater level and the red line shows the modelled seawater levels after the tidal signals have been filtered out on the left axis. The modelled seawater levels without the influence of mean wave period and direction (pink), without the influence of pressure (black) and without the influence of wind speed and direction (cyan) are presented as dashed lines. The atmospheric pressure is provided on the right axis.

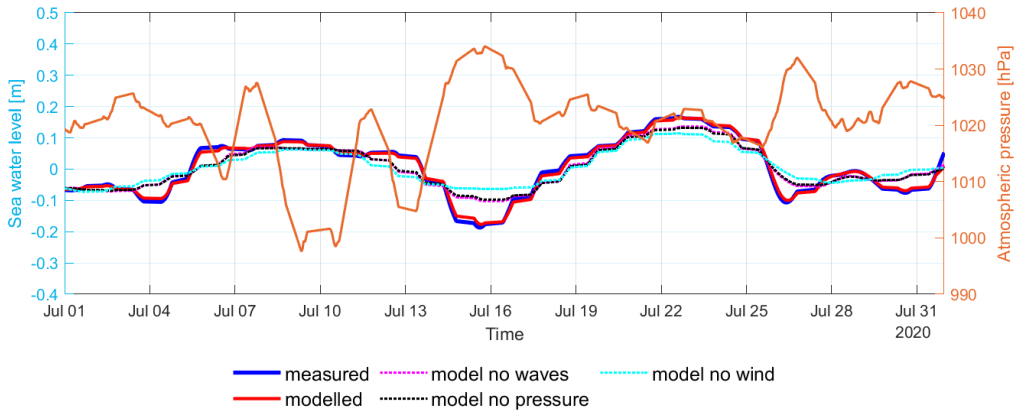


Figure 3.4: Final model validation for the storm that occurred in Simon's Town in July 2020. The blue line indicates the measured seawater level and the red line shows the modelled seawater levels after the tidal signals have been filtered out on the left axis. The modelled seawater levels without the influence of mean wave period and direction (pink), without the influence of pressure (black) and without the influence of wind speed and direction (cyan) are presented as dashed lines.The atmospheric pressure is provided on the right axis.

Figures 3.2–3.4 show that the NARX model is more accurate with all five meteocean variables (surface pressure, wind direction and speed, and mean wave period and direction) being fed as the inputs into the model. The R decreases and the RMSE values increase slightly without the presence of one or the other inputs, especially atmospheric pressure which coincides with the PCA results that this variable is one of the strongest influence on seawater level. Figures 3.2–3.4 visually show the decrease in accuracy (i.e. the model's inability to capture the peaks as seen across the validation period of all three storm events) of the model when the other input variables are omitted. Overall, the lowest R and highest RMSE values ranging from 0.85–0.95 and 18.26–100.5 mm respectively, are seen when the surface pressure is not included as an input into the NARX model. In June 2017, Simon's Town had lowest R and higher RMSE values compared to the other two sites (Durban and Port Elizabeth). It can also be seen that the Willmott index for this station and storm was the lowest when pressure and waves were not included in the model (0.44 and 0.17 respectively), this correlates with the high RMSE (136.4 and 100.5 mm) and bias values (718.7 and 707.3 mm). It can also be visually seen that the modelled seawater levels for Simon's Town failed to predict some of the peaks and example can be seen in

June 4 and 7, 2017, of Figure 3.2a). Modelled predictions for Durban and Port Elizabeth (June, 2017) had a R value of 0.99 and a Willmott index ranging from 0.97–0.99 when all metocean variables were inputted into the model. However, there is slight noise present in the modelled seawater level for Durban and there are areas where the seawater level was underestimated (May 8 to 10, 2007) and overestimated (March 3 and 20, 2007) (Figure 3.3).

Table 3.2: Correlation coefficients (R), the Willmott, et al., (2012) refined index of model performance, root mean square errors (RMSE), and bias for the three storm events.

Event	**Station**	**Variables**	***R***	**Willmott index**	**RMSE (mm)**	**Bias (mm)**
June 2017	Durban	Predicted	0.99	0.97	5.783	-19.74
		Model no waves	0.97	0.90	17.21	93.51
		Model no pressure	0.95	0.86	24.46	14.72
		Model no wind	0.97	0.90	17.09	83.26
	Port Elizabeth	Predicted	0.99	0.97	7.118	7.417
		Model no waves	0.99	0.95	11.46	44.69
		Model no pressure	0.95	0.86	33.18	196.5
		Model no wind	0.99	0.95	11.80	52.97
	Simon's Town	Predicted	0.97	0.63	60.78	288.8
		Model no waves	0.85	0.44	136.4	718.7
		Model no pressure	0.90	0.17	100.5	707.3
		Model no wind	0.96	0.79	41.91	277.3
March 2007	Durban	Predicted	0.99	0.99	5.680	7.464
		Model no waves	0.98	0.90	26.65	105.3
		Model no pressure	0.95	0.85	18.26	171.5
		Model no wind	0.97	0.90	18.25	104.8
July 2020	Simon's Town	Predicted	0.99	0.98	4.344	-6.129
		Model no waves	0.89	0.84	40.24	694.2
		Model no pressure	0.94	0.77	24.00	376.1
		Model no wind	0.94	0.84	29.01	685.3

The over and underestimated seawater level predictions that are seen in Figures 3.2–3.4 may be due to the oceanographic processes that were not accounted for in the development of the model, e.g. processes such as shoaling where the amplitude of the wave is increased as it runs from deep to shallower waters (Van Dongeren et al., 2007; Rautenbach et al., 2020). Large, wide continental shelves are commonly the most susceptible to experience shoaling as water piles up against the coast due to shallow seas (Rautenbach et al., 2020) and in turn are more vulnerable to storm surge events. This process occurs on the south coast at the Agulhas and Namaqua Bank (Rautenbach et al., 2020), where the continental shelf is wide and irregularly shaped (Carter er al., 1987). Since the

continental shelf on the east coast of South Africa is narrow (Schumann, 1987; Rautenbach et al.,2020), water is unable to pile up on easily on the shelf and thus storm surges are mainly caused by low atmospheric pressure (Rautenbach et al., 2020). Coastally trapped waves (CTWs) are an additional oceanographic process that occur at the south coast which are caused by the features of the continental shelf and coastline. The propagation of these waves depend on where they occur (on the western boundary of the oceans propagation is toward the equator and on the eastern boundary the wave propagation is toward the poles). CTWs are largely driven by the changes in the longshore component of the wind (Tomczak, 1998). The occurrence of CTWs on the south coast (Schumann and Brink, 1990) is also unaccounted for in the NARX models thus affecting the model performance.

The exclusion of variables that may have an influence on the coastal seawater level may also be a possibility of why the models underestimate the observed seawater levels. Storm surge is predominantly influenced by the atmospheric barometric pressure known as the inverted barometer effect, the set-up of the waves and wind as well as the bathymetry (Rautenbach et al., 2020). In this study, the bathymetry was not included as one of the variables that influenced seawater level predictions. In addition to bathymetry, friction also has a significant role on the control over tidal dynamics on the continental shelves and shallow waters. Bathymetry and the representation of friction can have a notable impact on the on coastal sea level (Wise et al., 2018; Serafin et al., 2019; Rasquin et al., 2020; Wachler et al., 2020).

In the Southern Hemisphere, the Agulhas Current is a strong, narrow (around 100 km wide) western boundary current which flows south along the east coast of South Africa, transporting warm, tropical water southward (Lutjeharms, 2006). The Agulhas Current receives its source from mesoscale eddies from the south of the Mozambique Channel and the East Madagascar Current (Morris and Lamont, 2019; Trott et al., 2021). In this study, the presence of the Agulhas Current was not taken into consideration when developing the models. This may

have a significant affect on the predictive skill of the NARX model on the stations along the south and east coast since the possible wave set-up caused by the Agulhas current system was not accounted for (Nhantumbo, 2019; Rautenbach et al., 2020) and may have been the cause of the noise present the modelled seawater level in Figure 3.3.

Chapter 4

Conclusions

The development of an open-loop NARX model using the Neural Net Time Series application from the Deep Learning Toolbox 14.0 provided by MATLAB ® (Mathworks, 2020) was created in order to predict seawater levels along the coast of South Africa. This research study provides an alternative method of predicting seawater levels along the coast of South Africa using data-driven models i.e. machine learning techniques in comparison to a recently developed numerical model which utilised the shallow water hydrodynamic model (Delft-3D) coupled with a non-stationary spectral wave model presented by Rautenbach et al. (2020).

The NARX model and the numerical model presented by Rautenbach et al. (2020) both aimed to develop a model that is able to predict storm surges along the coast of South Africa. Rautenbach et al. (2020) focused on one significant storm event where the model performed well for the stations situated on the West coast (Cape Town, Mossel Bay, and Port Elizabeth) of South Africa but decreased in model accuracy for the stations as we move along the south to the East coast of South Africa (East London, Durban, and Richard's Bay). Nonetheless, the results presented in Rautenbach et al. (2020) for the seawater level prediction accuracies were well within the limits of published literature. The NARX model utilised three storm events and produced results that are

more accurate than the numerical model for all three stations irrespective of their locations. Rautenbach et al. (2020) also included modelled wind speed and direction into their numerical model which can undoubtedly be explored in future research for the predictions of seawater level.

It has been demonstrated that the NARX model can provide valuable and more accurate information in comparison to the current numerical model in place by Rautenbach et al. (2020) which in turn can be applied to aid current forecasting systems in place at SAWS where marine weather forecasts are high-resolution but can only forecast for the next three days. The forecasting method provided in this study can have the potential to give high-resolution forecasts for more than three days at a time if it can be incorporated with the current systems in place.

The performance of the constructed NARX models showed that seawater levels can be predicted with high accuracy which indicated that the model could capture the non-linear relationships that exist between each forcing variable (Zubier and Eyouni, 2020). The model was also tested to investigate the role of each of the forcing metocean variables had on the seawater level and how the model performance differs if one variable was omitted. This revealed that each of the metocean variables played a significant role in the accuracy of the model's prediction since the model performance was most accurate in predicting seawater level when wind, wave, and atmospheric pressure were inputted into the network. A drastic decrease in the performance of the model is seen when the surface pressure is not included as an input into the NARX model (Table 3.2). The magnitude of how strong the influence of each metocean variable was also investigated using PCA where it demonstrated that surface pressure is the dominant driving force behind the seawater level (Figure 3.1 and Table 3.1).

The study presented here clearly shows an effective methodology to not only demonstrate the high accuracy it has on seawater level predictions, but also be able to further investigate the importance of what each oceanic and atmospheric variable has on the seawater level. Based on these conclusions, further investi-

gations can be done to explore other drivers that affect seawater level such as bathymetry and other oceanographic processes (e.g. shoaling). Future studies can involve understanding the local dynamics of the Agulhas Current and the effect that it may have on seawater level along the East and South coast of South Africa. PCA can also be performed once the influence of the Agulhas Current has been understood, to assess how significant the current is on seawater level compared to the metocean variables discussed in this study. Furthermore, research can be implemented to predict other oceanic and atmospheric parameters such as wind, waves, and atmospheric pressure. However, it must be noted that a mechanistic component along with the data assimilation may always be required to obtain the background signal. In this study, the component referred to is surface pressure since it has a strong influence on seawater level and is a significant input variable. This can lead to accurate predictions for other extreme events which use other forms of machine learning such as ARIMA, ARMA, and FFNN to predict waves, SST, and wind.

Bibliography

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M. and Ghemawat, S., 2016. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467.

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. and Kudlur, M., 2016. Tensorflow: A system for large-scale machine learning. In 12th USENIX symposium on operating systems design and implementation (OSDI 16) (pp. 265-283).

Abdi, H. and Williams, L.J., 2010. Principal component analysis. Wiley interdisciplinary reviews: computational statistics, 2(4), pp.433-459.

Agrawal, J.D. and Deo, M.C., 2002. On-line wave prediction. Marine structures, 15(1), pp.57-74.

Ahmed, A.N., Othman, F.B., Afan, H.A., Ibrahim, R.K., Fai, C.M., Hossain, M.S., Ehteram, M. and Elshafie, A., 2019. Machine learning methods for better water quality prediction. Journal of Hydrology, 578, p.124084.

Aït-Sahalia, Y. and Xiu, D., 2015. Using principal component analysis to estimate a high dimensional factor model with high-frequency data. Technical Report.

Akhtar, R. ed., 2019. Extreme Weather Events and Human Health: International Case Studies. Springer Nature.

Alaboud, T.M. and El-Bisy, M.S., 2011. Longshore current velocities prediction: using a neural networks approach. WIT Transactions on Ecology and the Environment, 149, pp.189-200.

Aljazeera.com. 2020. Storm Kills Several, Displaces Thousands In Cape Town. [online] Available at: https://www.aljazeera.com/news/2017/06/storm-kills-displaces-thousands-cape-town-170608052748704.html [Accessed 13 August 2020].

Alpaydin, E., 2020. Introduction to machine learning. MIT press.

Archetti, R., Paci, A., Carniel, S. and Bonaldo, D., 2016. Optimal index related to the shoreline dynamics during a storm: the case of Jesolo beach. Natural

Hazards and Earth System Sciences, 16(5), pp.1107-1122.

Artificial Neural Network for Precipitation and Water Level Predictions of Bedup River. IAENG International Journal of computer science, 34(2).l water level predictions. Ocean Engineering, 30(17), pp.2275-2295.

Artificial Neural Network for Precipitation and Water Level Predictions of Bedup River. IAENG International Journal of computer science, 34(2).l water level predictions. Ocean Engineering, 30(17), pp.2275-2295.

Balogun, A.L. and Adebisi, N., 2021. Sea level prediction using ARIMA, SVR and LSTM neural network: assessing the impact of ensemble Ocean-Atmospheric processes on models' accuracy. Geomatics, Natural Hazards and Risk, 12(1), pp.653-674.

Barbosa, S.M., Silva, M.E. and Fernandes, M.J., 2008. Changing seasonality in North Atlantic coastal sea level from the analysis of long tide gauge records. Tellus A: Dynamic Meteorology and Oceanography, 60(1), pp.165-177.

Beal, L.M., De Ruijter, W.P., Biastoch, A., Zahn, R., Cronin, M., Hermes, J., Lutjeharms, J., Quartly, G., Tozuka, T., Baker-Yeboah, S. and Bornman, T., 2011. On the role of the Agulhas system in ocean circulation and climate. Nature, 472(7344), p.429.

Bellprat, O. and Doblas-Reyes, F., 2016. Attribution of extreme weather and climate events overestimated by unreliable climate simulations. Geophysical Research Letters, 43(5), pp.2158-2164.

Beniston, M. and Stephenson, D.B., 2004. Extreme climatic events and their evolution under changing climatic conditions. Global and planetary change, 44(1-4), pp.1-9.

Bishop, C.M., 2006. Pattern recognition and machine learning. springer.

Bolton, J.J., Anderson, R.J., Smit, A.J. and Rothman, M.D., 2012. South African kelp moving eastwards: the discovery of Ecklonia maxima (Osbeck) Papenfuss at De Hoop Nature Reserve on the south coast of South Africa. African Journal of Marine Science, 34(1), pp.147-151.

Branch, G., 2017. Two oceans: a guide to the marine life of southern Africa. Penguin Random House South Africa.

Brander, R.W. and MacMahan, J.H., 2011. Future challenges for rip current research and outreach. Rip Currents, Beach Safety, Physical Oceanography and Wave Modeling, edited by: Leatherman, S. and Fletemeyer, J., CRC Press, Boca Raton, FL, pp.1-29.

Geach, B. (2020) Mean Monday: Another vicious winter storm surges through city of Cape Town. [online] Available at: https://www.dailymaverick.co.za/article/2020-07-13-mean-monday-another-vicious-winter-storm-surges-through-city-of-cape-town/ [Accessed 02 July 2021].

Caldwell, P. C., M. A. Merrifield, P. R. Thompson (2015), Sea level measured by tide gauges from global oceans — the Joint Archive for Sea Level holdings (NCEI Accession 0019568), Version 5.5, NOAA National Centers for Environmental Information, Dataset, doi:10.7289/V5V40S7W.

Caldwell, P. C., M. A. Merrifield, P. R. Thompson (2015), Sea level measured by tide gauges from global oceans — the Joint Archive for Sea Level holdings (NCEI Accession 0019568), Version 5.5, NOAA National Centers for Environmental Information, Dataset, doi:10.7289/V5V40S7W.

Cao, Q., Ewing, B.T. and Thompson, M.A., 2012. Forecasting wind speed with recurrent neural networks. European Journal of Operational Research, 221(1), pp.148-154.

Chambers, D.P., Cazenave, A., Champollion, N., Dieng, H., Llovel, W., Forsberg, R., von Schuckmann, K. and Wada, Y., 2017. Evaluation of the global mean sea level budget between 1993 and 2014. In Integrative Study of the Mean Sea Level and Its Components (pp. 315-333). Springer, Cham.

Chen, H. and Sun, J., 2015. Changes in climate extreme events in China associated with warming. International Journal of Climatology, 35(10), pp.2735-2751.

Chen, X., Zhang, X., Church, J.A., Watson, C.S., King, M.A., Monselesan, D., Legresy, B. and Harig, C., 2017. The increasing rate of global mean sea-level rise during 1993–2014. Nature Climate Change, 7(7), p.492.

Cid, A., Wahl, T., Chambers, D.P. and Muis, S., 2018. Storm surge reconstruction and return water level estimation in Southeast Asia for the 20th century. Journal of Geophysical Research: Oceans, 123(1), pp.437-451.

Davis-Reddy, C.L. and Vincent, K. 2017: Climate Risk and Vulnerability: A Handbook for Southern Africa (2nd Ed), CSIR, Pretoria, South Africa.

Defratti, M., Scholar, M. and Yarnal, B., 2015. Analyzing Fire Hazard Risk: A Case Study in Table Mountain National Park, Cape Town, South Africa. The Penn State McNair Journal, 20, pp.36-46.

Deo, M.C., Jha, A., Chaphekar, A.S. and Ravikant, K., 2001. Neural networks for wave forecasting. Ocean engineering, 28(7), pp.889-898.

Department of Environmental Affairs (2018). South Africa's Third National Communication under the United Nations Framework Convention on Climate Change. URL:https://unfccc.int/sites/default/files/resource/South%20African%20TNC%20Report%20%20to%20the%20UNFCCC_31%20Aug.pdf

Di Nunno, F. and Granata, F., 2020. Groundwater level prediction in Apulia region (Southern Italy) using NARX neural network. Environmental Research, 190, p.110062.

Di Piazza, A., Di Piazza, M.C. and Vitale, G., 2016. Solar and wind forecasting

by NARX neural networks. Renewable Energy and Environmental Sustainability, 1, p.39.

Dlugokencky, E. and Tans, P.: Trends in atmospheric carbon dioxide, National Oceanic & Atmospheric Administration, Earth System Research Laboratory (NOAA/ESRL), available at: http://www.esrl.noaa.gov/gmd/ccgg/trends/global.html, last access: 3 March 2019.

Donat, M.G., Lowry, A.L., Alexander, L.V., O'Gorman, P.A. and Maher, N., 2016. More extreme precipitation in the world's dry and wet regions. Nature Climate Change, 6(5), p.508.

Dudek, G., 2019. Generating random weights and biases in feedforward neural networks with random hidden nodes. Information sciences, 481, pp.33-56.

Easterling, D.R., Evans, J.L., Groisman, P.Y., Karl, T.R., Kunkel, K.E. and Ambenje, P., 2000. Observed variability and trends in extreme climate events: a brief review. Bulletin of the American Meteorological Society, 81(3), pp.417-426

Erfianto, B., Suwastika, N.A. and Prabowo, S., 2018, May. Decision System for Reservoir Upwelling Using Fuzzy Logic Based on Internet of Things. In 2018 6th International Conference on Information and Communication Technology (ICoICT) (pp. 370-375). IEEE.

Farber T., Hyman A. and Alexander E., 2020. [online] Ferocious storm blows into Cape Town. Available at: https://www.timeslive.co.za/news/south-africa/2020-07-13-watch-ferocious-storm-blows-into-cape-town/ [Accessed 06 July 2021].

Faris, H., Aljarah, I. and Mirjalili, S., 2016. Training feedforward neural networks using multi-verse optimizer for binary classification problems. Applied Intelligence, 45(2), pp.322-332.

Ferrario, F., Beck, M.W., Storlazzi, C.D., Micheli, F., Shepard, C.C. and Airoldi, L., 2014. The effectiveness of coral reefs for coastal hazard risk reduction and adaptation. Nature communications, 5, p.3794.

Filippo, A., Torres Jr, A.R., Kjerfve, B. and Monat, A., 2012. Application of Artificial Neural Network (ANN) to improve forecasting of sea level. Ocean & Coastal Management, 55, pp.101-110.

Frölicher, T.L. and Laufkötter, C., 2018. Emerging risks from marine heat waves. Nature communications, 9(1), p.650.

Frölicher, T.L., Fischer, E.M. and Gruber, N., 2018. Marine heatwaves under global warming. Nature, 560(7718), p.360.

Gardner, A.S., Moholdt, G., Cogley, J.G., Wouters, B., Arendt, A.A., Wahr, J., Berthier, E., Hock, R., Pfeffer, W.T., Kaser, G. and Ligtenberg, S.R., 2013. A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009. science, 340(6134), pp.852-857.

Garrabou, J., Coma, R., Bensoussan, N., Bally, M., Chevaldonné, P., Cigliano, M., Díaz, D., Harmelin, J.G., Gambi, M.C., Kersting, D.K. and Ledoux, J.B., 2009. Mass mortality in Northwestern Mediterranean rocky benthic communities: effects of the 2003 heat wave. Global change biology, 15(5), pp.1090-1103.

Ghorbani, M.A., Khatibi, R., Aytek, A., Makarynskyy, O. and Shiri, J., 2010. Sea water level forecasting using genetic programming and comparing the performance with artificial neural networks. Computers & Geosciences, 36(5), pp.620-627.

Grady, A.E., Moore, L.J., Storlazzi, C.D., Elias, E. and Reidenbach, M.A., 2013. The influence of sea level rise and changes in fringing reef morphology on gradients in alongshore sediment transport. Geophysical Research Letters, 40(12), pp.3096-3101.

Gupta, S. and Gupta, A., 2019. Dealing with noise problem in machine learning data-sets: A systematic review. Procedia Computer Science, 161, pp.466-474.

Han, G. and Shi, Y., 2008. Development of an Atlantic Canadian coastal water level neural network model. Journal of Atmospheric and Oceanic Technology, 25(11), pp.2117-2132.

Hansom, J.D., Switzer, A.D. and Pile, J., 2015. Extreme waves: Causes, characteristics, and impact on coastal environments and society. In Coastal and Marine Hazards, Risks, and Disasters (pp. 307-334). Elsevier.

Harris, D.L., 1962. The equivalence between certain statistical prediction methods and linearized dynamical methods. Mon. Wea. Rev, 90, pp.331-340.

Haupt, S.E., Pasini, A. and Marzban, C. eds., 2008. Artificial intelligence methods in the environmental sciences. Springer Science & Business Media. Herring, S.C., Hoerling, M.P., Peterson, T.C. and Stott, P.A., 2014. Explaining extreme events of 2013 from a climate perspective. Bulletin of the American Meteorological Society, 95(9), pp. S1-S104.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2018): ERA5 hourly data on single levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (Accessed on ¡ 22-DEC-2021 ¿), 10.24381/cds.adbb2d47

Higham, D.J. and Higham, N.J., 2016. MATLAB guide. Society for Industrial and Applied Mathematics.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H. and Kanae, S., 2013. Global flood risk under climate change. Nature Climate Change, 3(9), p.816.

Hobday, A.J., Alexander, L.V., Perkins, S.E., Smale, D.A., Straub, S.C., Oliver, E.C., Benthuysen, J.A., Burrows, M.T., Donat, M.G., Feng, M. and Holbrook, N.J., 2016. A hierarchical approach to defining marine heatwaves. Progress in

Oceanography, 141, pp.227- 238.

Huang, W., Murray, C., Kraus, N. and Rosati, J., 2003. Development of a regional neural network for coasta Bustami, R., Bessaih, N., Bong, C. and Suhaili, S., 2007.

Hutchings, L., Van der Lingen, C.D., Shannon, L.J., Crawford, R.J.M., Verheye, H.M.S., Bartholomae, C.H., Van der Plas, A.K., Louw, D., Kreiner, A., Ostrowski, M. and Fidel, Q., 2009. The Benguela Current: An ecosystem of four components. Progress in Oceanography, 83(1-4), pp.15-32.

IPCC, 2019: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Portner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegria, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]. In press.

IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. B. R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)]. Cambridge University Press. In Press.

Jury, M.R., 2013. Climate trends in southern Africa. South African Journal of Science, 109(1- 2), pp.1-11.

Jury, M.R., Valentine, H.R. and Lutjeharms, J.R., 1993. Influence of the Agulhas Current on summer rainfall along the southeast coast of South Africa. Journal of Applied Meteorology, 32(7), pp.1282-1287.

Kämpf, J. and Chapman, P., 2016. Upwelling systems of the world. Springer International Publishing Switzerland.

Karimi, S., Kisi, O., Shiri, J. and Makarynskyy, O., 2013. Neuro-fuzzy and neural network techniques for forecasting sea level in Darwin Harbor, Australia. Computers & Geosciences, 52, pp.50-59.

Kirezci, E., Young, I.R., Ranasinghe, R., Muis, S., Nicholls, R.J., Lincke, D. and Hinkel, J., 2020. Projections of global-scale extreme sea levels and resulting episodic coastal flooding over the 21st Century. Scientific reports, 10(1), pp.1-12.

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., Pickers, P. A., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Arneth, A., Arora, V. K., Barbero, L., Bastos, A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Doney, S. C., Gkritzalis, T., Goll, D. S., Harris, I., Haverd, V., Hoffman, F. M., Hoppema, M., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Johannessen, T., Jones, C. D., Kato, E., Keeling, R. F., Goldewijk, K. K., Landschützer, P., Lefèvre, N., Lienert, S., Liu, Z., Lombardozzi, D., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Olsen, A., Ono, T., Patra, P., Peregon, A., Peters, W., Peylin, P., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rocher, M., Rödenbeck, C., Schuster,

U., Schwinger, J., Séférian, R., Skjelvan, I., Steinhoff, T., Sutton, A., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., Viovy, N., Walker, A. P., Wiltshire, A. J., Wright, R., Zaehle, S., and Zheng, B.: Global Carbon Budget 2018, Earth Syst. Sci. Data, 10, 2141-2194, https://doi.org/10.5194/essd-10-2141-2018, 2018.

Lee, T.L., 2006. Neural network prediction of a storm surge. Ocean Engineering, 33(3-4), pp.483-494.

Lehmann, A. and Myrberg, K., 2008. Upwelling in the Baltic Sea—a review. Journal of Marine Systems, 74, pp. S3-S12.

Lemonnier, P. and Ainsworth, E.A., 2018. Crop Responses to Rising Atmospheric [CO2] and Global Climate Change. Food Security and Climate Change, pp.51-69.

Levitus, S., Antonov, J.I., Boyer, T.P., Baranova, O.K., Garcia, H.E., Locarnini, R.A., Mishonov, A.V., Reagan, J.R., Seidov, D., Yarosh, E.S. and Zweng, M.M., 2012. World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. Geophysical Research Letters, 39(10).

Li et al. (2012) integrated FL into their research which consisted of estimating the generations of internal waves in the northwestern parts of the Pacific Ocean. The results of the research were acceptable and it was concluded that FL systems may be used as a supplement for traditional methods such as neural networks.

Li, W., Wu, H., Zhu, N., Jiang, Y., Tan, J. and Guo, Y., 2020. Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit (GRU). Information Processing in Agriculture.

Lindsey, R. 2020. Climate Change: Atmospheric Carbon Dioxide. [online] Available at: https://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide [Accessed: 07 July 2021].

Lo-yat, A., Simpson, S.D., Meekan, M., Lecchini, D., Martinez, E. and Galzin, R., 2011. Extreme climatic events reduce ocean productivity and larval supply in a tropical reef ecosystem. Global Change Biology, 17(4), pp.1695-1702.

Lou, R., Wang, W., Li, X., Zheng, Y. and Lv, Z., 2021. Prediction of ocean wave height suitable for ship autopilot. IEEE Transactions on Intelligent Transportation Systems.

Lutjeharms, J.R., 2006. The agulhas current (Vol. 5). Berlin: Springer.

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

Makarynskyy, O., Pires-Silva, A.A., Makarynska, D. and Ventura-Soares, C.,

2005. Artificial neural networks in wave predictions at the west coast of Portugal. Computers & Geosciences, 31(4), pp.415-424.

Mandal, S. and Prabaharan, N., 2006. Ocean wave forecasting using recurrent neural networks. Ocean engineering, 33(10), pp.1401-1410.

Marine.weathersa.co.za. 2022. SAWS Marine - Storm Surge Forecasts. [online] Available at: https://marine.weathersa.co.za/Forecasts_Surge.html [Accessed 16 February 2022].

Maryan, C., Hoque, M.T., Michael, C., Ioup, E. and Abdelguerfi, M., 2019. Machine learning applications in detecting rip channels from images. Applied Soft Computing, 78, pp.84-93.

Mason, S.J., Waylen, P.R., Mimmack, G.M., Rajaratnam, B. and Harrison, J.M., 1999. Changes in extreme rainfall events in South Africa. Climatic Change, 41(2), pp.249-257.

Masselink, G., Castelle, B., Scott, T., Dodet, G., Suanez, S., Jackson, D. and Floc'h, F., 2016. Extreme wave activity during 2013/2014 winter and morphological impacts along the Atlantic coast of Europe. Geophysical Research Letters, 43(5), pp.2135-2143.

Mentaschi, L., Vousdoukas, M.I., Voukouvalas, E., Dosio, A. and Feyen, L., 2017. Global changes of extreme coastal wave energy fluxes triggered by intensified teleconnection patterns. Geophysical Research Letters, 44(5), pp.2416-2426.

Merrifield, M.A., Becker, J.M., Ford, M. and Yao, Y., 2014. Observations and estimates of wave-driven water level extremes at the Marshall Islands. Geophysical Research Letters, 41(20), pp.7245-7253.

Mirza, M.M.Q., 2003. Climate change and extreme weather events: can developing countries adapt? Climate policy, 3(3), pp.233-248.

Monserrat, S. and Rabinovich, A.B., 2006. Meteotsunamis: atmospherically induced destructive ocean waves in the tsunami frequency band. Natural Hazards and Earth System Science, 6(6), pp.1035-1051.

Montana, D.J. and Davis, L., 1989, August. Training feedforward neural networks using genetic algorithms. In IJCAI (Vol. 89, pp. 762-767).

Mohri, M., Rostamizadeh, A. and Talwalkar, A., 2018. Foundations of machine learning. MIT press.

More, A. and Deo, M.C., 2003. Forecasting wind with neural networks. Marine structures, 16(1), pp.35-49.

Morris, T. and Lamont, T., 2019. Using ocean robots on high-resolution profiling to capture the fast-flowing Agulhas Current. South African Journal of Science, 115(1-2), pp.1-3.

Muis, S., Apecechea, M.I., Dullaart, J., de Lima Rego, J., Madsen, K.S., Su, J., Yan, K. and Verlaan, M., 2020. A High-Resolution Global Dataset of Extreme Sea Levels, Tides, and Storm Surges, Including Future Projections. Frontiers in Marine Science, 7, p.263.

Muslim, T.O., Ahmed, A.N., Malek, M.A., Abdulmohsin Afan, H., Khaleel Ibrahim, R., El-Shafie, A., Sapitang, M., Sherif, M., Sefelnasr, A. and El-Shafie, A., 2020. Investigating the influence of meteorological parameters on the accuracy of sea-level prediction models in Sabah, Malaysia. Sustainability, 12(3), p.1193.

Nascimento, S. and Franco, P., 2009, September. Segmentation of upwelling regions in sea surface temperature images via unsupervised fuzzy clustering. In International Conference on Intelligent Data Engineering and Automated Learning (pp. 543-553). Springer, Berlin, Heidelberg.

Nerem, R.S., Beckley, B.D., Fasullo, J.T., Hamlington, B.D., Masters, D. and Mitchum, G.T., 2018. Climate-change–driven accelerated sea-level rise detected in the altimeter era. Proceedings of the National Academy of Sciences, 115(9), pp.2022-2025.

Nhantumbo, B. J. (2019). Drivers of coastal sea level variability along the east and south 817 of South Africa. University of Cape Town.Retrieved from https://open.uct.ac.za/handle/11427/30348.

Nitsure, S.P., Londhe, S.N. and Khare, K.C., 2014. Prediction of sea water levels using wind information and soft computing techniques. Applied Ocean Research, 47, pp.344-351.

Nizamani, Z., Yih, L.W., Wahab, M.M.A. and Mustaffa, Z., 2017. Determination of Correlation for Extreme Metocean Variables. In MATEC Web of Conferences (Vol. 103, p. 04013). EDP Sciences.

O'Gorman, P.A., 2015. Precipitation extremes under climate change. Current climate change reports, 1(2), pp.49-59.

Oliver, E.C., Benthuysen, J.A., Darmaraki, S., Donat, M.G., Hobday, A.J., Holbrook, N.J., Schlegel, R.W. and Sen Gupta, A., 2021. Marine heatwaves. Annual review of marine science, 13, pp.313-342.

Pashova, L. and Popova, S., 2011. Daily sea level forecast at tide gauge Burgas, Bulgaria using artificial neural networks. Journal of Sea Research, 66(2), pp.154-161.

Patil, K. and Deo, M.C., 2017. Prediction of daily sea surface temperature using efficient neural networks. Ocean Dynamics, 67(3-4), pp.357-368.

Patil, K. and Deo, M.C., 2018. Basin-Scale Prediction of Sea Surface Temperature with Artificial Neural Networks. Journal of Atmospheric and Oceanic Technology, 35(7), pp.1441- 1455.

Prein, A.F., Rasmussen, R.M., Ikeda, K., Liu, C., Clark, M.P. and Holland, G.J., 2017. The future intensification of hourly precipitation extremes. Nature Climate Change, 7(1), p.48.

Proshutinsky, A., Ashik, I.M., Dvorkin, E.N., Häkkinen, S., Krishfield, R.A. and Peltier, W.R., 2004. Secular sea level change in the Russian sector of the Arctic Ocean. Journal of Geophysical Research: Oceans, 109(C3).

Pugh, D., 1987. Tides, surges and mean sea level. A handbook for engineers and scientists. Quataert, E., Storlazzi, C., Rooijen, A., Cheriton, O. and Dongeren, A., 2015. The influence of coral reefs and climate change on wave-driven flooding of tropical coastlines. Geophysical Research Letters, 42(15), pp.6407-6415.

Rasquin, C., Seiffert, R., Wachler, B. and Winkel, N., 2020. The significance of coastal bathymetry representation for modelling the tidal response to mean sea level rise in the German Bight. Ocean Science, 16(1), pp.31-44.

Ray, G.C., 1991. Coastal-zone biodiversity patterns. Bioscience, 41(7), pp.490-498.

Rautenbach, C., Daniels, T., Vos, M. and Barnes, M.A., 2020. A coupled wave, tide and storm surge operational forecasting system for South Africa: validation and physical description. Natural Hazards: Journal of the International Society for the Prevention and Mitigation of Natural Hazards, pp.1-33.

Reason, C.J.C., 2001. Evidence for the influence of the Agulhas Current on regional atmospheric circulation patterns. Journal of Climate, 14(12), pp.2769-2778.

Rodríguez-Ramírez, A., Pérez-Asensio, J.N., Santos, A., Jiménez-Moreno, G., Villarías- Robles, J.J., Mayoral, E., Celestino-Pérez, S., Cerrillo-Cuenca, E., López-Sáez, J.A., León, Á. and Contreras, C., 2015. Atlantic extreme wave events during the last four millennia in the Guadalquivir estuary, SW Spain. Quaternary Research, 83(1), pp.24-40.

Rosenzweig, C., Iglesias, A., Yang, X.B., Epstein, P.R. and Chivian, E., 2001. Climate change and extreme weather events; implications for food production, plant diseases, and pests. Global change & human health, 2(2), pp.90-104.

Rothman, M.D., Anderson, R.J. and Smit, A.J., 2006. The effects of harvesting of the South African kelp (Ecklonia maxima) on kelp population structure, growth rate and recruitment. In Eighteenth International Seaweed Symposium (pp. 109-115). Springer, Dordrecht.

Rouault, M., Penven, P. and Pohl, B., 2009. Warming in the Agulhas Current system since the 1980's. Geophysical Research Letters, 36(12).

Rouault, M., Roy, S.S. and Balling, R.C., 2013. The diurnal cycle of rainfall in South Africa in the austral summer. International Journal of Climatology, 33(3), pp.770-777.

Rouault, M., White, S.A., Reason, C.J.C., Lutjeharms, J.R.E. and Jobard, I., 2002. Ocean– atmosphere interaction in the Agulhas Current region and a South African extreme weather event. Weather and Forecasting, 17(4), pp.655-669.

Sahoo, B. and Bhaskaran, P.K., 2019. Prediction of storm surge and coastal inundation using Artificial Neural Network–A case study for 1999 Odisha Super Cyclone. Weather and Climate Extremes, p.100196.

Schumann, E.H., 1987. The coastal ocean off the east coast of South Africa. Transactions of the Royal Society of South Africa, 46(3), pp.215-229.

Schumann, E.H. and Brink, K.H., 1990. Coastal-trapped waves off the coast of South Africa: generation, propagation and current structures. Journal of Physical Oceanography, 20(8), pp.1206-1218.

Sembiring, L., Van Ormondt, M., Van Dongeren, A. and Roelvink, D., 2015. A validation of an operational wave and surge prediction system for the Dutch coast. Natural Hazards & Earth System Sciences Discussions, 3(6).

Serafin, K.A., Ruggiero, P., Barnard, P.L. and Stockdon, H.F., 2019. The influence of shelf bathymetry and beach topography on extreme total water levels: Linking large-scale changes of the wave climate to local coastal hazards. Coastal Engineering, 150, pp.1-17.

Shamshirband, S., Mosavi, A., Rabczuk, T., Nabipour, N. and Chau, K.W., 2020. Prediction of significant wave height; comparison between nested grid numerical model, and machine learning models of artificial neural networks, extreme learning and support vector machines. Engineering Applications of Computational Fluid Mechanics, 14(1), pp.805-817.

Shi, E., Li, Q., Gu, D. and Zhao, Z., 2018, February. A method of weather radar echo extrapolation based on convolutional neural networks. In International Conference on Multimedia Modeling (pp. 16-28). Springer, Cham

Smale, D.A. and Wernberg, T., 2013. Extreme climatic event drives range contraction of a habitat-forming species. Proceedings of the Royal Society B: Biological Sciences, 280(1754), p.20122829

Smith, A.M., Guastella, L.A., Bundy, S.C. and Mather, A.A., 2007. Combined marine storm and Saros spring high tide erosion events along the KwaZulu-Natal coast in March 2007. South African Journal of Science, 103(7-8), pp.274-276.

Soman, S.S., Zareipour, H., Malik, O. and Mandal, P., 2010, September. A review of wind power and wind speed forecasting methods with different time horizons. In North American Power Symposium 2010 (pp. 1-8). IEEE.

Song, D.S. and Bae, H.K., 2011. Observation and forecasting of rip current generation in Haeundae Beach, Korea Plan and Experiment. Journal of Coastal Research, pp.946-950.

Sousa, F.M., Nascimento, S., Casimiro, H. and Boutov, D., 2008. Identification of upwelling areas on sea surface temperature images using fuzzy clustering. Remote Sensing of Environment, 112(6), pp.2817-2823.

Stocker, T. ed., 2014. Climate change 2013: the physical science basis: Working Group I contribution to the Fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Stott, P., 2016. How climate change affects extreme weather events. Science, 352(6293), pp.1517-1518.

Stott, P.A., Christidis, N., Otto, F.E., Sun, Y., Vanderlinden, J.P., Van Oldenborgh, G.J., Vautard, R., Von Storch, H., Walton, P., Yiou, P. and Zwiers, F.W., 2016. Attribution of extreme weather and climate-related events. Wiley Interdisciplinary Reviews: Climate Change, 7(1), pp.23-41.

Tadesse, M., Wahl, T. and Cid, A., 2020. Data-driven modeling of global storm surges. Frontiers in Marine Science, p.260.

Tapoglou, E., Forster, R.M., Dorrell, R.M. and Parsons, D., 2021. Machine learning for satellite-based sea-state prediction in an offshore windfarm. Ocean Engineering, 235, p.109280.

Climate Risk Profile: South Africa (2021): The World Bank Group.

Tomczak, M., 1998. Shelf and coastal oceanography. University of Maine.

Trauth, M.H., Gebbers, R., Marwan, N. and Sillmann, E., 2007. MATLAB recipes for earth sciences (Vol. 34). Berlin: Springer.

Trenberth, K.E., Fasullo, J.T. and Shepherd, T.G., 2015. Attribution of climate extreme events. Nature Climate Change, 5(8), p.725.

Tripathi, K.C., Das, I.M.L. and Sahai, A.K., 2006. Predictability of sea surface temperature anomalies in the Indian Ocean using artificial neural networks.

Troell, M., Robertson-Andersson, D., Anderson, R.J., Bolton, J.J., Maneveldt, G., Halling, C. and Probyn, T., 2006. Abalone farming in South Africa: an overview with perspectives on kelp resources, abalone feed, potential for on-farm seaweed production and socio-economic importance. Aquaculture, 257(1-4), pp.266-281.

Trott, C.B., Subrahmanyam, B. and Washburn, C.E., 2021. Investigating the response of temperature and salinity in the agulhas current region to enso events. Remote Sensing, 13(9), p.1829.

Tsai, C.P., Lin, C. and Shen, J.N., 2002. Neural network for wave forecasting among multi-stations. Ocean engineering, 29(13), pp.1683-1695.

Tyson, P.D. and Preston-Whyte, R.A., 2000. Weather and climate of southern Africa. Oxford University Press.

Vabalas, A., Gowen, E., Poliakoff, E. and Casson, A.J., 2019. Machine learning algorithm validation with a limited sample size. PloS one, 14(11), p.e0224365.

Valentine, H.R., Lutjeharms, J.R.E. and Brundrit, G.B., 1993. The water masses and volumetry of the southern Agulhas Current region. Deep Sea Research Part I: Oceanographic Research Papers, 40(6), pp.1285-1305.

Van der Lingen CD. Diet of sardine Sardinops sagax in the southern Benguela upwelling ecosystem. South African Journal of Marine Science. 2002 Jun 1;24(1):301-16.

Van Dongeren, A.R.J.A., Battjes, J., Janssen, T., Van Noorloos, J., Steenhauer, K., Steenbergen, G. and Reniers, A.J.H.M., 2007. Shoaling and shoreline dissipation of low-frequency waves. Journal of Geophysical Research: Oceans, 112(C2).

Velimirov, B., Field, J.G., Griffiths, C.L. and Zoutendyk, P., 1977. The ecology of kelp bed communities in the Benguela upwelling system. Helgoländer wissenschaftliche Meeresuntersuchungen, 30(1), p.495.

Verheye, H.M., Lamont, T., Huggett, J.A., Kreiner, A. and Hampton, I., 2016. Plankton productivity of the benguela current large marine ecosystem (bclme). Environmental Development, 17, pp.75-92.

Vigaud, N., Richard, Y., Rouault, M. and Fauchereau, N., 2009. Moisture transport between the South Atlantic Ocean and southern Africa: relationships with summer rainfall and associated dynamics. Climate Dynamics, 32(1), pp.113-123.

Vilibić, I. and Šepić, J., 2009. Destructive meteotsunamis along the eastern Adriatic coast: Overview. Physics and Chemistry of the Earth, Parts A/B/C, 34(17-18), pp.904-917.

Wang, G., Wang, D., Trenberth, K.E., Erfanian, A., Yu, M., Bosilovich, M.G. and Parr, D.T., 2017. The peak structure and future changes of the relationships between extreme precipitation and temperature. Nature Climate Change, 7(4), p.268.

Wachler, B., Seiffert, R., Rasquin, C. and Kösters, F., 2020. Tidal response to sea level rise and bathymetric changes in the German Wadden Sea. Ocean Dynamics, pp.1-20.

Wang, X., Guo, P. and Huang, X., 2011. A review of wind power forecasting models. Energy procedia, 12, pp.770-778.

Weisberg, R.H. and Zheng, L., 2006. Hurricane storm surge simulations for Tampa Bay. Estuaries and Coasts, 29(6), pp.899-913.

Wernberg, T., 2021. Marine heatwave drives collapse of kelp forests in Western Australia. In Ecosystem collapse and climate change (pp. 325-343). Springer, Cham.

Wernberg, T., Smale, D.A., Tuya, F., Thomsen, M.S., Langlois, T.J., De Bettignies, T., Bennett, S. and Rousseaux, C.S., 2013. An extreme climatic event alters marine ecosystem structure in a global biodiversity hotspot. Nature Climate Change, 3(1), p.78.

Winter, G., Hetzel, Y., Huang, P., Hipsey, M.R., Mulligan, R.P. and Hansen, J., 2019. 3 Coastal processes, extreme events and forecasting. Marine Extremes: Ocean Safety, Marine Health and the Blue Economy

Wise, A., Hughes, C.W. and Polton, J.A., 2018. Bathymetric influence on the coastal sea level response to ocean gyres at western boundaries. Journal of Physical Oceanography, 48(12), pp.2949-2964.

Willmott, C. J., Robeson, S. M., and, Matsuura, K. (2012). A refined index of model performance. International Journal of Climatology, 32(13), 2088–2094. https://doi.org/10.1002/joc.2419

Wright, L.D., Resio, D.T. and Nichols, C.R., 2019. Causes and Impacts of Coastal Inundation. In Tomorrow's Coasts: Complex and Impermanent (pp. 103-118). Springer, Cham.

Wu, A., Hsieh, W.W. and Tang, B., 2006. Neural network forecasts of the tropical Pacific sea surface temperatures. Neural Networks, 19(2), pp.145-154.

Xu, X., Dong, X., Lin, W., Yun, R. and Zhu, D., 2018, July. Wind field retrieving for SCAT onboard CFOSAT based on PCA method. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium (pp. 5587-5590). IEEE.

Xu, A., Chang, H., Xu, Y., Li, R., Li, X. and Zhao, Y., 2021. Applying artificial neural networks (ANNs) to solve solid waste-related issues: A critical review. Waste Management, 124, pp.385-402.

Yu, Z., Niu, Z., Tang, W. and Wu, Q., 2019. Deep learning for daily peak load forecasting–a novel gated recurrent neural network combining dynamic time warping. IEEE access, 7, pp.17184-17194.

Zubier, K.M. and Eyouni, L.S., 2020. Investigating the role of atmospheric variables on sea level variations in the Eastern Central Red Sea using an artificial neural network approach. Oceanologia, 62(3), pp.267-290.