# Prediction of Impending Drought Scenarios Based on Surface and Subsurface Parameters in a Selected Region of Tropical Queensland, Australia

**Bithin Datta[1]\*, Dilip Kumar Roy[2], Jonathan J. Kelley[1], Bradley Stevens[1]**

[1]Discipline of Civil Engineering, College of Science and Engineering, James Cook University, Townsville, Australia
[2]Irrigation and Water Management Division, Bangladesh Agricultural Research Institute, Joydebpur, Bangladesh
Email: \*bithin.datta@jcu.edu.au, dilip.roy@my.jcu.edu.au

## Abstract

Droughts occur in all climatic regions around the world costing a large expense to global economies. Reasonably accurate prediction of drought events helps water managers proper planning for utilization of limited water resources and distribution of available waters to different sectors and avoid catastrophic consequences. Therefore, a means to create a simplistic approach for forecasting drought conditions with easily accessible parameters is highly desirable. This study proposes and evaluates newly developed accurate prediction models utilizing various hydrologic, meteorological, and geohydrology parameters along with the use of Artificial Neural Network (ANN) models with various forecast lead times. The present study develops a multitude of forecasting models to predict drought indices such as the Standard Precipitation Index with a lead-time of up to 6 months, and the Soil Moisture Index with a lead-time of 3 months. Furthermore, prediction models with the capability of approximating surface and groundwater storage levels including the Ross River Dam level have been developed with relatively high accuracy with a lead-time of 3 months. The results obtained from these models were compared to current values, revealing that ANN based approach can be used as a simple and effective predictive model that can be utilized for prediction of different aspects of drought scenarios in a typical study area like Townsville, North Queensland, Australia which had suffered severe recent drought conditions for almost six recent years (2014 to early 2019).

## Keywords

Drought Prediction, Artificial Neural Network, Prediction Model, Dam Levels, Aquifer Salinity, Water Storage, Townsville Queensland

## 1. Introduction

The main problem with mitigating the consequences of a drought is that a drought is generally identified only after the drought has started and we are well into a drought scenario. This lack of proper early prediction of droughts can prevent efficient management of drought without adequate foresight. The tropical northern part of the Queensland state in Australia had been reeling under a severe drought for several years. Although drought conditions in Townsville and surrounding areas were declared officially in 2015, the drought had been in full swing by then with visible signs of extreme water scarcity. Because of the severe constraints on water use, as the result of the existing drought condition, the economy of the region suffered substantially. Early detection and prediction of the drought could have lessened the enormous economic costs and the associated hardships. Therefore, this study was undertaken to develop at least preliminary models capable of predicting future drought scenarios in this region with some forecast lead time. Such predictions can enormously help manage the supply and demand sides of water resources, with enormous benefits resulting from early detection and prediction of droughts. Both subsurface storage levels as well as surface storage in the Ross River dam situated in Townsville, Northern Queensland were incorporated in the prediction models for drought scenarios. In addition, aquifer salinity level prediction models were developed.

The term drought has been recognized around the world as an environmental disaster costing a large expense to global economies. The occurrences of droughts are likely in most climatic regions around the world following rainfall deficiencies over an extended period. The recent droughts in the state of California in the U.S.A. is a glaring example. Temperatures, high winds, low relative humidity, timing, and characteristics of rains including distribution of rainy days during crop growing seasons as well as onset, intensity, duration, and terminations of rain play a significant role in the occurrence of drought [1]. Droughts can effectively be classified into three main categories: meteorological drought, agricultural drought, and hydrological drought [2]. Meteorological drought is considered as the ongoing shortage of precipitation, which usually affects a large area (*i.e.*, North Queensland). Meteorological drought is the originating source of drought that triggers the other two types of droughts. Agricultural drought is related to the scarcity of water needed to meet the requirement of crop production or plant growth. Hydrological drought can be defined as the occurrence of low flow in watercourses, and the below average levels of lakes and groundwater. It is related to a period with a decrease in surface and groundwater resources available for uses of an established water resource system [1]. A lack of precipitation often triggers agricultural and hydrological droughts, but other factors, including more intense but less frequent precipitation, poor water management, and erosion, can also cause or enhance these droughts.

Among others, lack of rain is the main cause of drought. It has, however, been noted that both Hydro-meteorological variables and climate indices have a sig-

nificant effect on the cause, duration, and intensity of a drought. However, the intensity of a drought may not solely depend on the rainfall occurrence [3]. The occurrence of precipitation is perhaps the fundamental parameter regarding the initiation of drought. As stated previously, meteorological drought occurs because of rainfall deficiencies that are ultimately a precursor for agricultural and hydrological drought. Therefore, the justification behind rainfall deficiencies is an important aspect of drought. The lack of rain can be linked to many causes related to climate indices and anomalies (such as the El Niño Southern Oscillation (ENSO) event) and possibly even climate change [4]. Whilst these climate indices and anomalies have a large influence on the occurrence of drought, present study does not address these issues, as the intention is to develop a simple prediction model within the scope of the data availability for the study area. Poor water management can often be one of the lead causes behind drought severity and localized occurrence. The use of ineffective and failed water management plans can effectively increase the severity of localized drought conditions with the inappropriate use of water.

The severity of droughts can be characterized by drought indexes that have been developed to detect, monitor, and assess drought events [5]. Drought indicators essentially are variables that identify and assess drought conditions, which are commonly based on hydrological and meteorological variables. There are numerous indices used throughout the industry that has played a major role in drought management. The most commonly used indices include Palmer Drought Severity Index (PDSI), Standard Precipitation Index (SPI), Soil Moisture Index (SMI), and Southern Oscillation Index (SOI). Drought indicators facilitate an effective way of understanding the characteristics of a particular drought event. Each of the above indicators is utilized for identification and classification of the drought event occurrence. Each of these indicators is dependent on the quantifying parameters used. For example, the SPI considers only precipitation, whilst PDSI considers precipitation, evapotranspiration, and runoff. Although the PDSI may be more reliable, SPI is a simplistic model that provides accurate analysis. The use of any of these indicators depends on the input data available and the region being considered (when calculating climate Indices), and these indicators provide useful stepping-stone to a forecasting model.

There are a few researchers, who have developed forecasting models to assist in the prediction of drought occurrence, intensity, and duration [6] [7] [8] [9]. These forecast models rely heavily on large amounts of spatial data and complex modelling procedures. Therefore, accurate and long-term data are needed to work effectively with these models. However, in many parts of the world, such as regional North Queensland, Australia has extremely limited amount of temporal and spatial data. As a result, many of the forecasting models that have been used elsewhere may not be suitable for the tropical Townsville region, in north Queensland, Australia due to the limited amount of relevant data. Therefore, it is necessary to construct modelling solutions that can derive desired output variables with minimal input variables. This may be achieved through modelling

networks such as Artificial Neural Network (ANN) models through which effective drought prediction, and management of its consequences can be achieved. If an effective drought prediction model can be derived, the management of water resources can reduce the implications of drought. Restriction schemes and water penalties can be applied in an effective manner and adequate distribution to agricultural regions may be achieved. Hence this attempt at preliminary development of drought models for this region. The recent severe drought in the Townsville and surrounding areas of tropical Northern Queensland resulting in severe constraints on availability and use of water was the motivation to try and develop drought prediction models which can be beneficial for more efficient management of water supply based on better forecasting of drought conditions. The ANN was chosen as a training and testing-based pattern recognition tool to learn and extract information from the patterns of input parameter values and the resulting drought conditions as outputs. It was also envisaged that the ANN based predictions also can utilize the capabilities of the ANN to recognize relevant input data contributing to better prediction. The different modelling tools used for prediction of drought scenarios at different parts of the world are briefly discussed here.

The probability of occurrence and the severity of drought can be forecasted by regression analysis [10]. The development of a regression analysis model that exhibited the grain yield of a crop was developed by Kumar and Panu [6]. The use of crop yield as an agricultural drought quantifying parameter and multiple variables affecting crop yield as explanatory variables was used for the assessment of agricultural drought forecasting [1]. The model generated by Kumar and Panu [6] is capable of explaining approximately 82 percent of the variation of grain yield and was able to predict the yield several months in advance. However, they found that the moisture index of the soil had marginal contribution to the analysis and accounted for insignificant variation in yield [6]. Another study conducted by Liu and Juárez [8] used ENSO indices and satellite recorded Normalized Difference Vegetation Index (NDVI) to construct a drought prediction model using a multiple linear regression technique for a region in Brazil. Three regression models were generated using an NDVI anomaly as a dependent variable and various ENSO index anomalies as independent variables. It was concluded in this study that the use of satellite recorded NDVI rather than precipitation data improved the correlation with the ENSO indices. Whereas, the third model generated, using a dataset with high anomaly values of NVDI and ENSO indices, predicted drought onset in the Brazilian region four months in advance with a success rate of 68 percent [8].

Time series analysis has also successfully been utilized for drought forecasting by several researchers in different parts of the world. The advantages of using this tool include its efficient search capability for estimations, identification, and analytical checking, capabilities important for model development [9]. Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) are two of the commonly used stochastic

models that are generally fitted to Time Series data. Furthermore, seasonal data can be utilized by using the Seasonal ARIMA model, which is written as, ARIMA $(p, d, q)(P, D, Q)_m$, in which the second part considers the seasonal data [7]. These models can be used through computer programming where auto-generation of the model is achieved by input data. This modelling system has been used in various drought applications including the modelling of a drought forecasting parameter. The use of both yearly and monthly PDI, utilized as stochastic input data for quantifying a parameter, was used in combination with ARIMA models to simulate and forecast PDI models by Rao and Padmanabhan [11]. Furthermore, a study conducted by Fernandez, *et al.* [12] using a multiplicative seasonal ARIMA model, had the ability to forecast stream flow with a lead-time of 12 months. Drought indices were derived based on this predicted average stream flow over different periods. However, the capability of this forecast is strongly dependent on the identification, estimation, and diagnostic checks to select the best model [10].

ANN models, mimicking the operation of a human brain can be utilized to forecast the values of the predicted set based on a set of predictors or input variables [13]. A major advantage of using an ANN model, especially for forecasting, is that intermediate relationships between inputs and outputs do not need to be fully defined. This capability of ANN makes it suitable for prediction of drought forecasting in which the relationship between the large numbers of complex input variables can produce suitable output solutions [10]. A good example of the application of ANN modelling approach in drought prediction is the work conducted by Morid, *et al.* [14], where they predicted quantitative values of drought indices. This involved using multiple combinations of rainfall data, SPI, and the Effective Drought Index (EDI) in the prior months. It was found that the best ANN, for both SPI and EDI, had quite a simple framework, where a three-layer network with a maximum of six neurons for a hidden layer was sufficient to find the optimal outputs. Furthermore, Kuligowski and Barros [13] developed a site-specific short-term precipitation-forecasting model based on ANN. The parameters utilized in this model consisted of radiosonde-based 700-hPa wind direction and precursor precipitation data to generate the forecast. By using these parameters, the ANN provided quantitative precipitation forecasts with a lead-time of up to 6 hours in advance. The justification behind using a 6-hour lead time is because the relationship between the precipitation quantity at any two locations significantly weaken as lag time increases when, precipitation fields change rapidly with time [13]. However, the ability to increase the lag time depends on additional parameters to be considered. The models to be generated however, will consider an array of different parameters where relationships between variables may be able to increase lead-time and forecast a much wider range of essential drought parameters and thresholds. The involvement of drought indices utilized not just as inputs into the model but also as an output will not only help predict droughts, but drought severity as well.

The aim of this paper is to develop and evaluate ANN based drought-forecasting

model by utilizing various hydrologic, meteorological, and geohydrology parameters. This model will be evaluated for possible application to data obtained from the North Queensland region to mitigate critical impacts on water scarcity. Several models will be generated to assess how different parameters may be affecting drought severity of the study area. Upon the successful evaluation of the model capabilities and accuracy, performance of the models is evaluated and tested with past/present data to ensure accuracy.

## 2. Methodology

### 2.1. Drought Indicators

Categorizing and assessing a drought is essential to determine when drought response actions need to be implemented. The use of drought indicators and triggers is important to assess the onset of drought conditions, monitor, and measure drought events, and to decrease drought impacts [15]. Drought indicators, essentially, are variables that identify and assess drought conditions, which are commonly based on hydrological and meteorological variables. Examples of such commonly used indicators include precipitation, soil moisture, stream flow, groundwater levels, surface water reservoir levels, temperature, and evapotranspiration. As an indicator breaches a certain threshold (often referred to as a trigger), this categorizes the drought and typically may depict levels of severity. The severity of droughts can be characterized by drought indexes that have been developed to detect, monitor, and assess drought events [5]. There are numerous indices used throughout the industry, which have played a major role in drought management where the most commonly used indices include PDSI, SPI, SMI, and ENSO. Brief description for each of these indices are provided in the following paragraphs.

The PDSI has been used for the past 50 years as a means of providing an individual measure of meteorological drought severity [16]. Three indicators are considered in the calculation of PDSI: Precipitation data, temperature data, and the available water content in the soil. From these indicators, variables such as soil recharge, evapotranspiration, runoff, and the moisture loss in the top soil layers can be found. The severity of the drought in question is considered to be a function of both magnitude and frequency of the moisture deficit [17]. However, there are a few limitations associated with the PDSI approach. These limitations include: an intrinsic time scale that makes it more suitable for agricultural applications, and not so much for hydrologic applications, that runoff only occurs after all soil layers are saturated causing underestimation, and the reaction time to respond to forming and diminishing droughts [1].

The SPI was developed to provide a better representation of the wetness and dryness in comparison to the PDSI approach. SPI was developed to quantify a precipitation deficit for different time scales [17]. It is basically, a standardizing transform of the probability of long-term precipitation that has been recorded [16]. This long-term record is then fitted to a probability distribution, which is transformed into a normal distribution so that the mean SPI for the location and

desired period is zero with a standard deviation of one [1]. Ideally, twenty to thirty years of monthly values (or more) would be optimal for accurate results. The time series of the SPI is then used for drought monitoring by setting specific thresholds of the SPI for defining the drought beginning and ending times. An accumulation of the SPI values can be used to assess the drought severity by using the indicators specified in [17]. There are, however, limitations that the SPI encounter where the length of rainfall data used has a significant impact on the SPI values. Mishra and Singh [1] suggested that when calculating SPI values, computed over different lengths of time, it may result in significantly inconsistent values. However, SPI is relatively simple to find given minimal inputs.

Soil moisture is an important parameter when considering the hydrological balance and surface interactions. However, none of the previously mentioned indices directly considers soil moisture content as part of the index. Therefore, the development of SMI was developed by the High Plains Regional Climate Centre to assist in the onset of agricultural drought by indicating the observed moisture of the soil relative to a plants ability to extract water [18]. The values obtained through calculations procedures form a distribution where the $\theta$ values are scaled from -5 (severely deficient water content) to 5 (very wet). This index has been utilized in many different scenarios since its development for calculating the water stress in Nebraska and has proven to be particularly useful [19]. El Niño and La Niña events occur naturally in the global climate system. The occurrences of these events are due to the change in state of the Pacific Ocean and the atmosphere surrounding it from its normal state for a period of several seasons. The effect of El Niño events in the Australian region (Western Pacific) associates with the cooling of the ocean, whereas La Niña events have the reverse effect. This however acts in reverse order when considering the central and eastern pacific region where the theory of these events was first acknowledged. The changes that continuously occur in the Pacific Ocean and its surrounding atmosphere occur in a cycle called the ENSO.

All the previously mentioned indices, and majority of others that have been derived, use precipitation as a primary parameter for the prediction and assessment of drought [1]. The cause of torrential or lack of precipitation has been proven to be related to atmospheric phenomena, for instance ENSO, Sea Surface Temperature (SST), and Geopotential Height (GpH) [10]. However, the relationship between the climate indices and precipitation is dependent on the region to be forecasted. For example, the use of ENSO data will be reliable in the pacific region such as Eastern Australia and Peru but, will not be a reliable forecasting tool for a region like Europe. In order to produce accurate drought-forecast models, the prediction of ENSO actions is important particularly in Australia. Three approaches are generally used in predicting ENSO:

1) Statistical Models (Oceanic & Atmospheric Models): currently the most used model used. However, is limited due to the lack of time scale data as records only go back to the 1950's.

2) Physical Ocean & Statistical Atmosphere Models: this method is the second

most used method. The forecasting of ENSO relies on the physical laws that govern the ocean's response to variations in atmospheric boundary conditions.

3) Physical Coupled Ocean & Atmosphere Models

These three schemes have forecast lead times that are typically similar whereas the simple models, 1 and 2, have an approximate accuracy for lead times to 1 year. Scenario 3 can predict slightly further in time, however, is a slightly more complex method [20]. These models provide a strong forecasting analysis tool (when used correctly) which has been utilized by many, particularly in Australia. These indices are great tools for drought prediction. However, they require a large scale of indicators to be able to predict rainfall data accurately.

## 2.2. Drought Forecasting

The forecasting of drought essentially requires past data (hydro-meteorological variables), a set of drought and climatic indices, and a suitable prediction model to predict the probability of occurrence as well as the nature of severity of drought [10]. The input variables for drought prediction vary depending on different types of droughts to be forecasted. For instance, historical precipitation data is needed to predict meteorological drought events. On the other hand, prediction of hydrological drought scenario necessarily demands reservoir water level, stream flow, and ground water levels data. Likewise, agricultural drought is predicted by using data on soil moisture content and crop yield data.

Several drought indices, discussed earlier, have been derived to assess the effect of a drought and defining different drought parameters. In addition, the use of climate indices such as ENSO is used in conjunction with the hydrological values for long-term drought forecasting [10]. The prediction of drought can be made by utilizing regression models, time series models, probability models, neural network models, and hybrid models [10]. Current study utilizes the implicit pattern recognition capability of a trained and tested ANN to develop a drought prediction model.

## 2.3. The Study Area and Data

The data used in this study are obtained from two sources: Bureau of Meteorology, Australia and the Townsville City Council, Queensland, Australia. The data on soil moisture, precipitation, evapotranspiration, temperature, and the historic Ross River dam (located in the study area) levels are collected from these two sources. The SPI and SMI indices are calculated using the procedures mentioned in the previous subsections of the methodology section. In addition, groundwater levels and groundwater salinity data are obtained from two bores within the Townsville catchment as shown in Figure 1 below.

The implementation of the usable datasets varied depending on the number of inputs and outputs to be specified, and whether the datasets are time dependent or not. The data used was organized and manipulated that includes shifting data to different prediction ranges e.g., for a 3-month prediction model, the target data was shifted up three-time steps with respect to the input data.
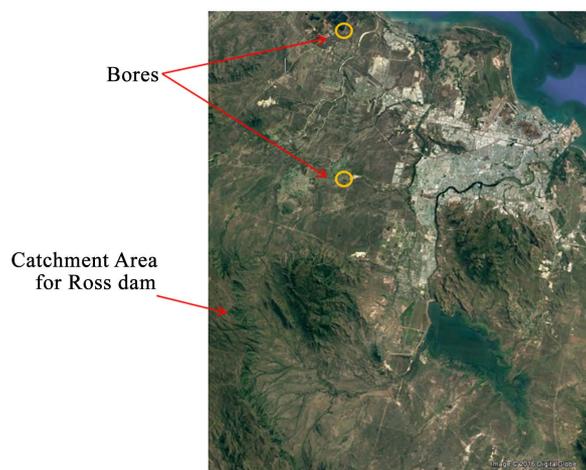
**Figure 1.** Study area (Townsville catchment with bore locations).

## 2.4. Model Selection/Training

This study utilizes time series neural networks as a prediction tool for future scenarios of drought forecasting. In particular, Non-Linear Autoregressive with External Input (NARX) model was utilized as the preferred neural network tool. This model uses both time series and regression analysis within its process to maximize the computational ability and efficiency to achieve optimal output results.

The number of layers utilized throughout testing was maintained at 2 layers: one hidden layer and one output layer with a consistent number of neurons (default of 10 used). Furthermore, a time delay of two-time steps was initiated to ensure the system can learn and start accurately approximating a value by the third time step in the training process as seen in **Figure 2**.

The NARX model takes the form of:

$$y(t) = f\left(x(t-1)\cdots x(t-d), y(t-1)\cdots y(t-d)\right) \quad (1)$$

where, $d$ = number of delays, and $t$ = time step.

The training performance is evaluated by computing the Mean Square Error (MSE) as shown in **Figure 3**, and repeatedly training the data until sufficient performance is identified.

When the inputs are not assumed to be time dependent, such as the salinity in bores, a different ANN structure, Feed Forward Back propagation (FFB) was used. This ANN structure does not feed the output data back as the next input like the NARX, rather each input is independent of one another, and the relationship is purely between the input and output without respect to time. The structure of the ANN model used for this purpose can be viewed in **Figure 4**. The training procedure is similar to the NARX model and follows the same performance steps.

After satisfactory training of the models, the trained and tested models are used for future predictions. The predicted results were compared directly against the actual values to validate the accuracy of the model and the simulation.
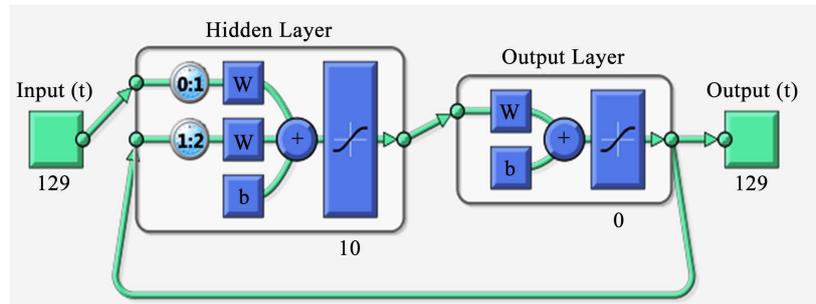
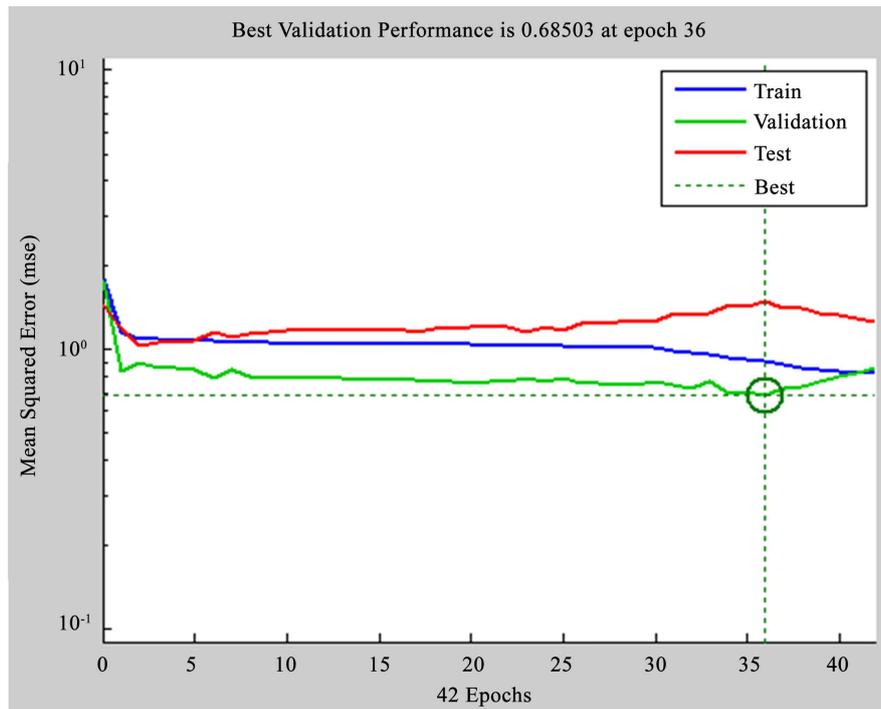**Figure 2.** Algorithm/ANN Map dependent on time.
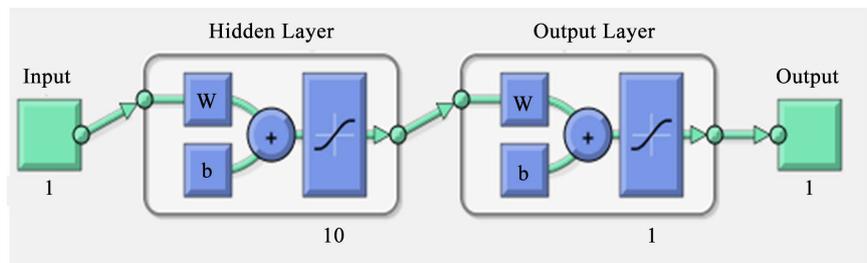


**Figure 3.** MSE plot.



**Figure 4.** Algorithm/ANN Map independent of time.

Hundreds of these tests were conducted for several scenarios where inputs and outputs were changed, length of prediction was altered, and input simplicity. The tests that acquired reasonable results were sub divided into three drought scenarios. Simulation can then be conducted with additional input variables once validation of results showed sufficient accuracy. This can be done by using

just one-time step at a time if required. The model will then produce a predicted output value for a specific lead-time for which the data was trained *i.e.*, three months.

## 3. Developed Models

A large number of ANN models were generated in the testing phase. However, the final and informative tests have been collaborated and identified under each drought type subgroup. Each model underwent the methodology previously stated above.

### 3.1. Meteorological Drought

A large number of ANN models were generated in the testing phase. However, the final and informative tests have been collated and identified under each drought type subgroup. Each model underwent the methodology previously stated above.

Prediction model 2 takes advantage of the SOI as well as the addition of monthly precipitation values and monthly average maximum temperature measurements. The purpose of this model is to include additional dynamic inputs to assist with a prolonged lead-time of 6 months.

Figure 5 represents the inputs and outputs to be used to generate a particular prediction model to find SPI values at a lead-time of 3 and 6 months. The model uses both agricultural and meteorological input parameters to assess the appropriate wetness intensity accurately. The purpose of interchanging the lead-time is to assess the error variation with a prolonged lead-time.

### 3.2. Hydrological Drought

There are two types of hydrological models generated for hydrological drought: Surface water levels, and ground water levels. The first model incorporates the use of drought indices as inputs to predict the water levels of the Ross River dam, as illustrated in Figure 6(a).
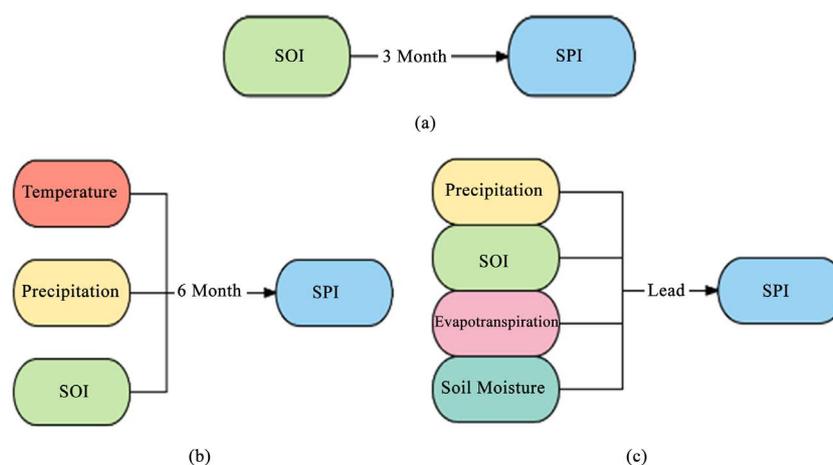


**Figure 5.** Developed models: (a) Model 1; (b) Model 2; (c) Models 3 and 4.
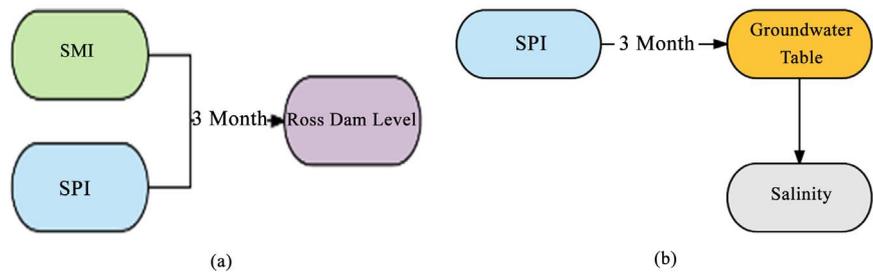
**Figure 6.** Developed models: (a) Model 5; (b) Models 6, 7, 8, and 9.

Two bores in Townsville, north Queensland, Australia were assessed in this next model, Bore 11800007A and Bore 11700027B. Firstly, the groundwater table was predicted using a NARX model with a lead-time of 3 months. Once reasonable results were obtained, a FFB model was used to link the predicted groundwater levels with the salinity of the ground water. The schematic diagram of the modelling approach is presented in **Figure 6(b)**. The purpose of using two different bores is to assess the conditions of a coastal aquifer in the vicinity of the coast, and also, relatively away from the coast. Then distinguish the different effecst the inputs have on the prediction.

### 3.3. Agricultural Drought

To assess agricultural drought conditions, the use of the SMI at the root zone soil is a good tool and indication of drought severity. Therefore, two prediction models were generated with the use of different input variables. The first prediction, as seen in **Figure 7(a)**, utilized the meteorological parameter SPI as an input. This parameter was used as there may be a delayed (3 Month) correlation between the wetness intensity, and the soil moisture at the root zone layer. By using a single variable, it makes the manipulation of data much simpler.

The final test utilizes the average monthly soil moisture and the total monthly evapotranspiration occurring. These parameters were chosen as input as they represent the movement of water in the soil itself. A maximum lead-time of 3 months was used, as this model does not account for any meteorological variables.

### 4. Results and Discussions

This section addresses the results obtained through testing, where the overall performance of the models was validated. Through these tests, a functional and simple ANN model was developed to predict meteorological, hydrological, and agricultural drought for the Townsville region, in Northern Queensland, Australia.

### 4.1. Meteorological Drought

For the assessment of meteorological drought, numerous tests were conducted with variations in input parameters with respect to meteorological drought
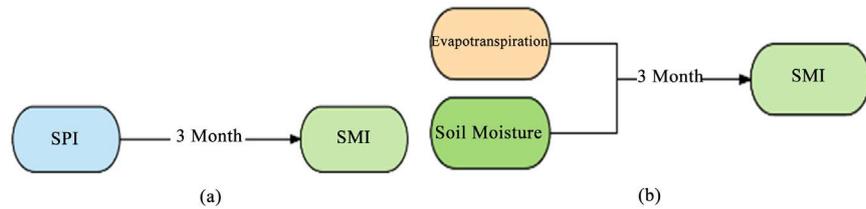
**Figure 7.** Developed models: (a) Model 10; (b) Model 11.

indices. With the available data, the obtained relationship between the inputs and accurate precipitation measures was not realistic for an acceptable lead-time. Therefore, present study used appropriate drought indices for the assessment of drought conditions of the specified study area. Moreover, the use of a drought index such as SPI drastically increased the ANN model ability to predict drought conditions for an appropriate lead-time. Four tests were conducted to assess meteorological drought prediction. The results of these tests are discussed here.

The first prediction model (Model 1) establishes the relationship between SOI and SPI. As seen from the trend between actual and predicted values in Figure 8(a), three months lead-time prediction of SPI from SOI is quite accurate since this model only assesses one variable, and the intensity of SPI is governed by many meteorological factors. The simplicity behind the use of one variable makes this prediction model a viable monthly model to grasp an understanding of the approximate wetness in three months. The accuracy of this model, however, can be explained by the effect of ENSO on the east coast of Australia. Generally, the more intense this phenomenon is, the greater the intensity of wetness in the region. As such, this model may be applied for prediction of drought scenarios on the Eastern coast of Australia by utilizing the limited available data. Considering one variable is used for this prediction, the model has an approximate accuracy of 43 percent, which is quite reasonable for one input variable.

The second model (Model 2) utilizes average monthly high temperature (with respect to daily maximum temperatures), monthly precipitation values, and SOI as inputs to approximate SPI. This model incorporates two additional inputs on top of prediction model1 to help gain a more accurate relationship over a longer 6-month ahead prediction. As seen in Figure 8(b), the predicted values follow an accurate trend with respect to actual SPI values. However, the amplitude of these prediction estimates is substantially lower, particularly when the SPI reaches the wetter stages (>0). This may be because, dry periods are magnified under certain temperatures; however, under wet conditions temperature may not have a substantial effect. This is most likely because the temperature is taken inland (Townsville) as opposed to out at sea, where increased heat enhances evaporation and in turn, increases precipitation and wetness. Therefore, the use of temperature and precipitation being used simultaneously may provide some sort of cancellation effect when considering a wet period. Despite this, the model prediction accuracy increases when considering intense drought scenario's (SPI < 0). An error analysis conducted on the dataset to assess the accuracy of the

(a)



(b)

(c)



(d)
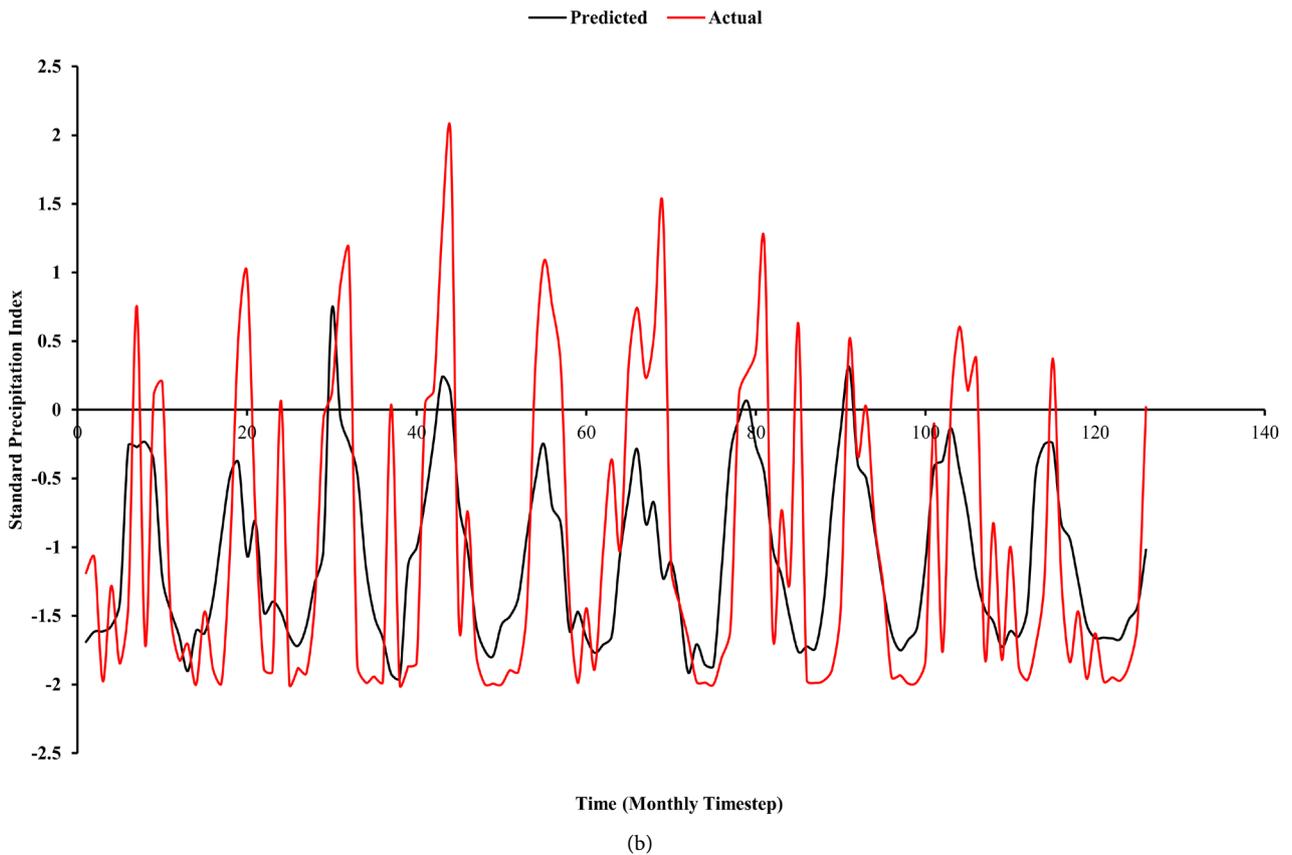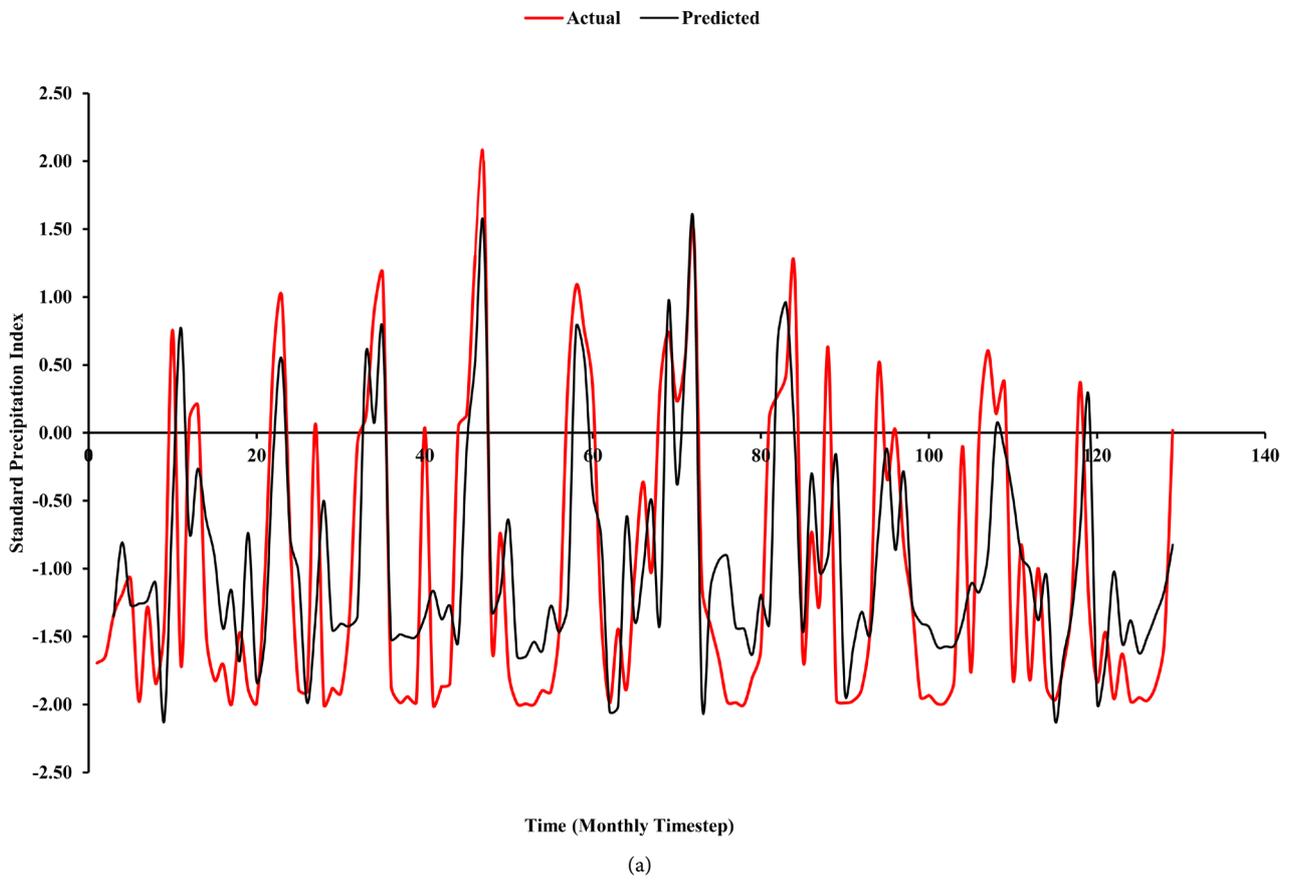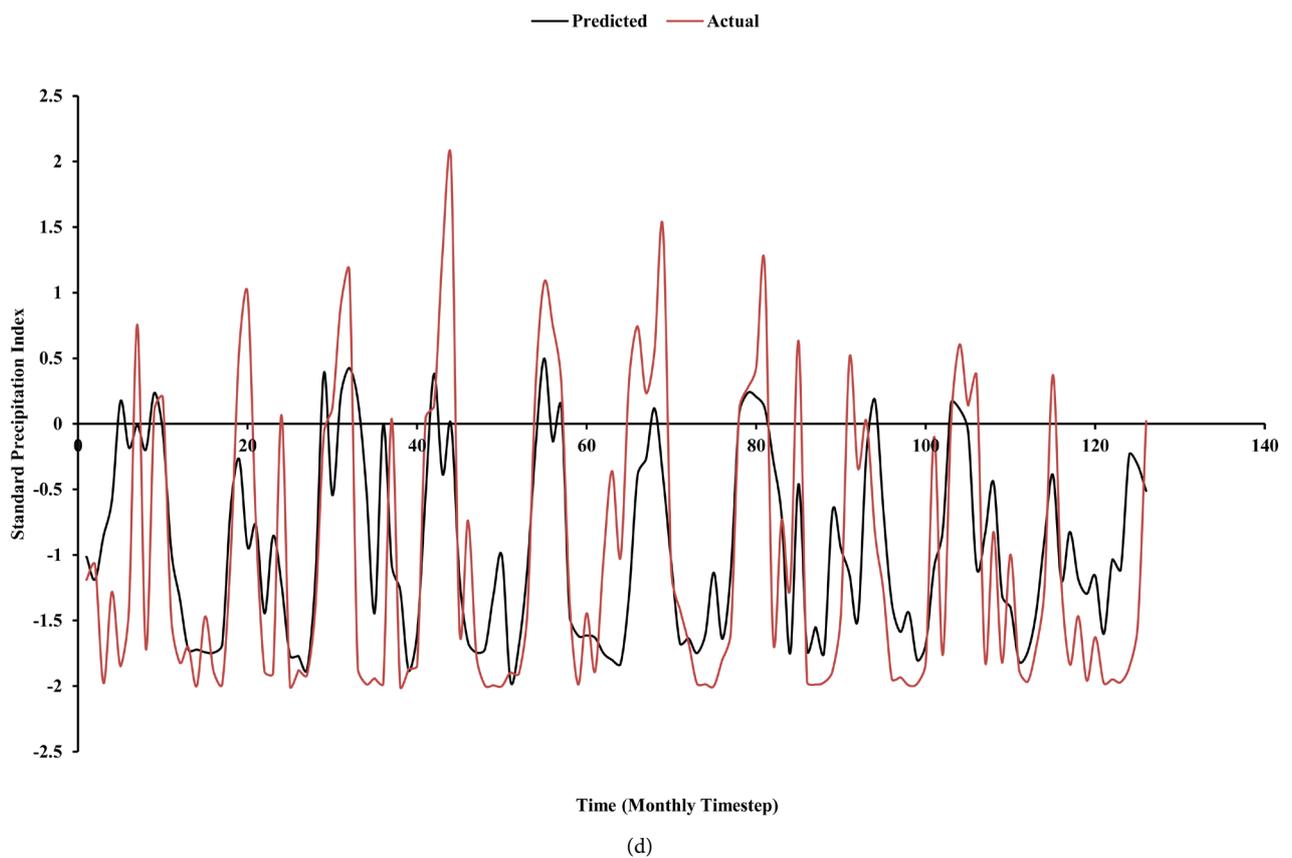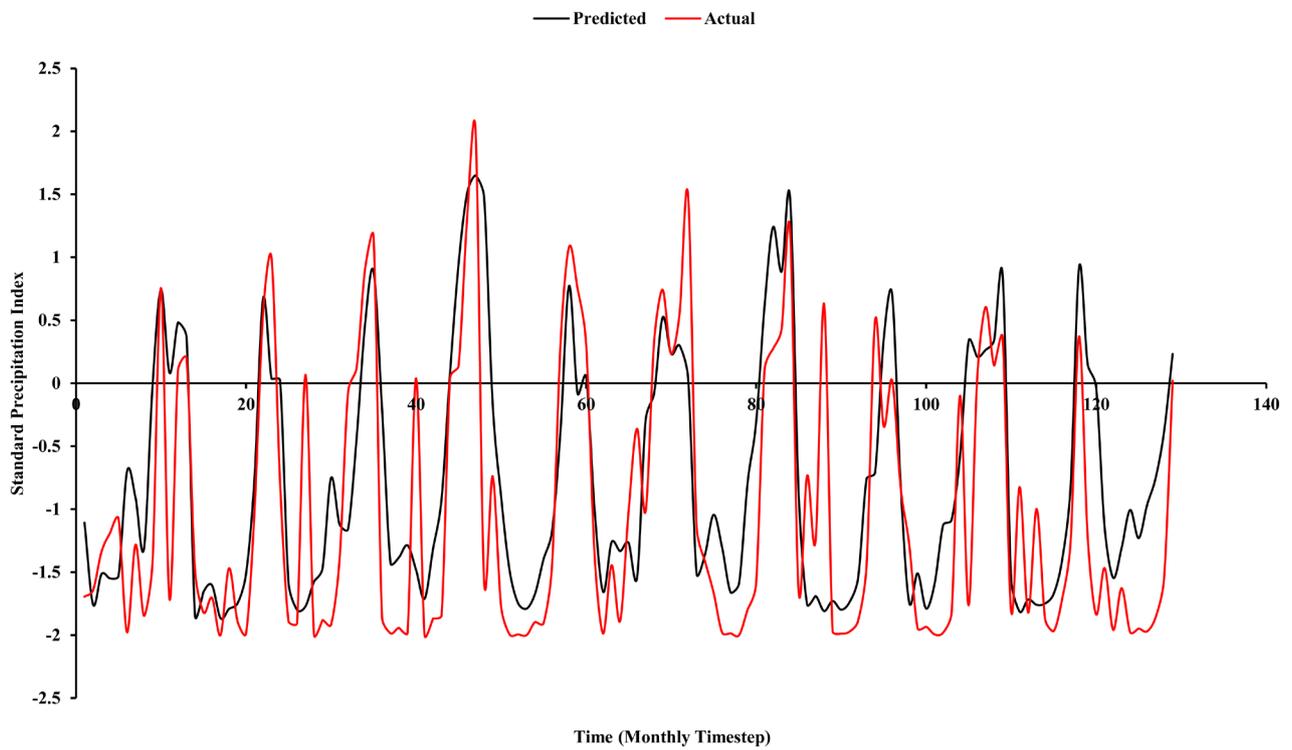
**Figure 8.** Meteorological drought prediction; (a) Model 1; (b) Model 2; (c) Model3 (3-month lead time); (d) Model 4 (6-month lead time).

model indicates that the average percentage error occurring is 76.4%, therefore which may seem quite low accuracy. It can be seen that the vast outliers will occur in the wetter periods in which the model does not quite meet the intensity of the SPI, consequently affecting the model performance. As previously stated, this model is more focused on the drought conditions and with that in mind, the error for drought condition assessment is substantially lower.

The third model (model 3) developed utilizes a wide range of input variables, namely monthly precipitation, SOI, monthly evapotranspiration, and average soil moisture. These parameters represent a greater refinement of the cause of drought conditions. Due to the inclusion of these additional inputs, the model provided better results than expected over the original three-month lead-time. Therefore, an additional model with a lead-time extended to six months was also implemented. As seen in Figure 8(c), there is a good trend matching between inputs and outputs, with matching amplitude indicating that there exists strong relationship between these datasets. Change in SOI generally enhances intensity of precipitation values, where precipitation values are also majorly governed by evapotranspiration, and hence the change in water content in the soil. With such relationships between inputs, a closer and more accurate model has been created that has more accurate prediction capabilities of future values. Error analysis indicates a lower error occurring when comparing with the previous model. With a percentage error of approximately 42 percent, this model can be labelled to be 58 percent accurate. The reduction in error in this model is generally due to the accuracy in intensity, particularly in the wetter periods. By utilizing the time series tool and reducing the overall lead-time to 3 months, the time steps can adapt with the changes in input variables more simply.

Furthermore, model providing six-month ahead prediction (model 4) as seen in Figure 8(d) also demonstrates a good trend of prediction. However, the intensity of the predicted SPI does not match perfectly with the actual values. This may simply be due to the prolonged lead-time of 6 months, hence reducing the effectiveness of the relationship between the input values. Moreover, it can be identified that the values react more appropriately with each given time step when comparing to model 2. Appropriate error calculations seem to provide some justification behind the model's improved ability with respect to model 2. Given an average accuracy of 34.7 percent, this model is a relatively accurate model especially considering the lead-time being 6 months.

## 4.2. Hydrological Drought

This section in particular concerns one of the more important issues surrounding urban developments and the agricultural industry. Five models have been developed for this purpose (Models 5 - 9). The results obtained from these models are presented and discussed in the following sub-sections.

### Model 5 (Surface Water: Ross River Dam Water Levels)

This prediction model incorporates the use of the two drought indices, SPI and SMI, to help predict Ross dam water levels. The use of additional inputs was

trialed prior to final validation of these inputs, however, did not improve neither accuracy nor simplicity of the model and output values. As can be seen in Figure 9(a), accurate interpretations of dam level readings with a lead-time of 3 months can effectively be found. The model generates an accurate trend that follows actual water levels which continues along time steps. However, in months of major water influx, the predicted dam levels typically lag one-time step behind which is expected whilst using drought indices as an input. This is not necessarily an overly disastrous occurrence as the model adequately mimics negative change in levels quite well. Therefore, the model is near perfect for drought mitigation and management strategy development. However, it is worth mentioning that the results do have fluctuations where predicted dataset tends to be slightly higher than actual values with slight fluctuations. A possible explanation for this is the exact location in which the inputs were taken. The data obtained is measured in the Townsville CBD, which is a substantial distance from both the Ross River dam and the catchment area itself. So, just because there was substantial rain in the CBD hence increasing the SPI and SMI thresholds, does not mean that the rainfall intensity and therefore drought indices will be identical within the dam's catchment. The resolution to this problem is to simply retrieve precipitation data within the catchment zone, if available, which could potentially reduce the error drastically and possibly increase the lead-time to something more respectable. Furthermore, this model does not consider water usage from external sources such as town water and the agricultural industry. Typically, with respect to pastures, crops, and other vegetation, the use of water peaks during the drier periods. This is due to the reduction of rainfall and therefore creating increasing demand of water to maintain vegetation life sufficiently. This may be another reason as to the decrease in accuracy as the dam levels decrease. The error presented in this model is substantially lower than previous models generated. At an accuracy of approximately 85 percent, it can be utilized as an efficient and effective tool to help manage water supply.

### Model 6 (Ground Water: Groundwater Table) Bore: 11700027B

This model assesses the effect and correlation between SPI and groundwater table on a coastal aquifer at bushland beach of Townsville, Queensland, Australia. The data in Figure 9(b) represents the depth measured from ground level to the datum level of the water and plotted against the predicted data. The selection of SPI without the inclusion of SMI was found to have the closest correlation. The results in Figure 9(b) show a good correlation between input and output values and show a clear response to changes in SPI values as seen for the first 10-time steps. Following this, the prediction overcompensates for several time steps and assumes a larger amount ground water is present. A possible scenario for this occurrence may be due to an increased ratio of volume with respect to the aquifer height, which the model may not have had enough experience within the training state. A solution to this would simply be to have the ability to access a larger dataset for a coastal aquifer in the region, which is currently not available. However, this model still proves to be accurate with minor error variations.
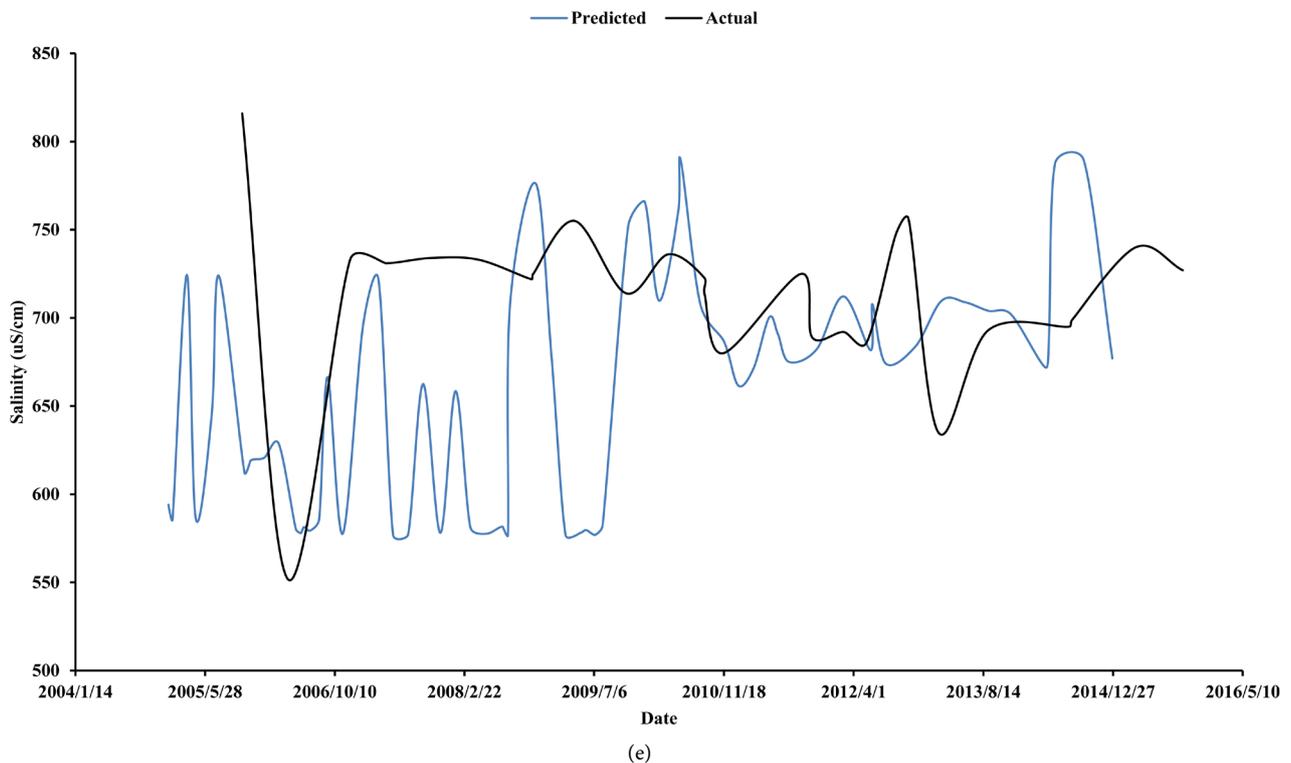
(a)



(b)
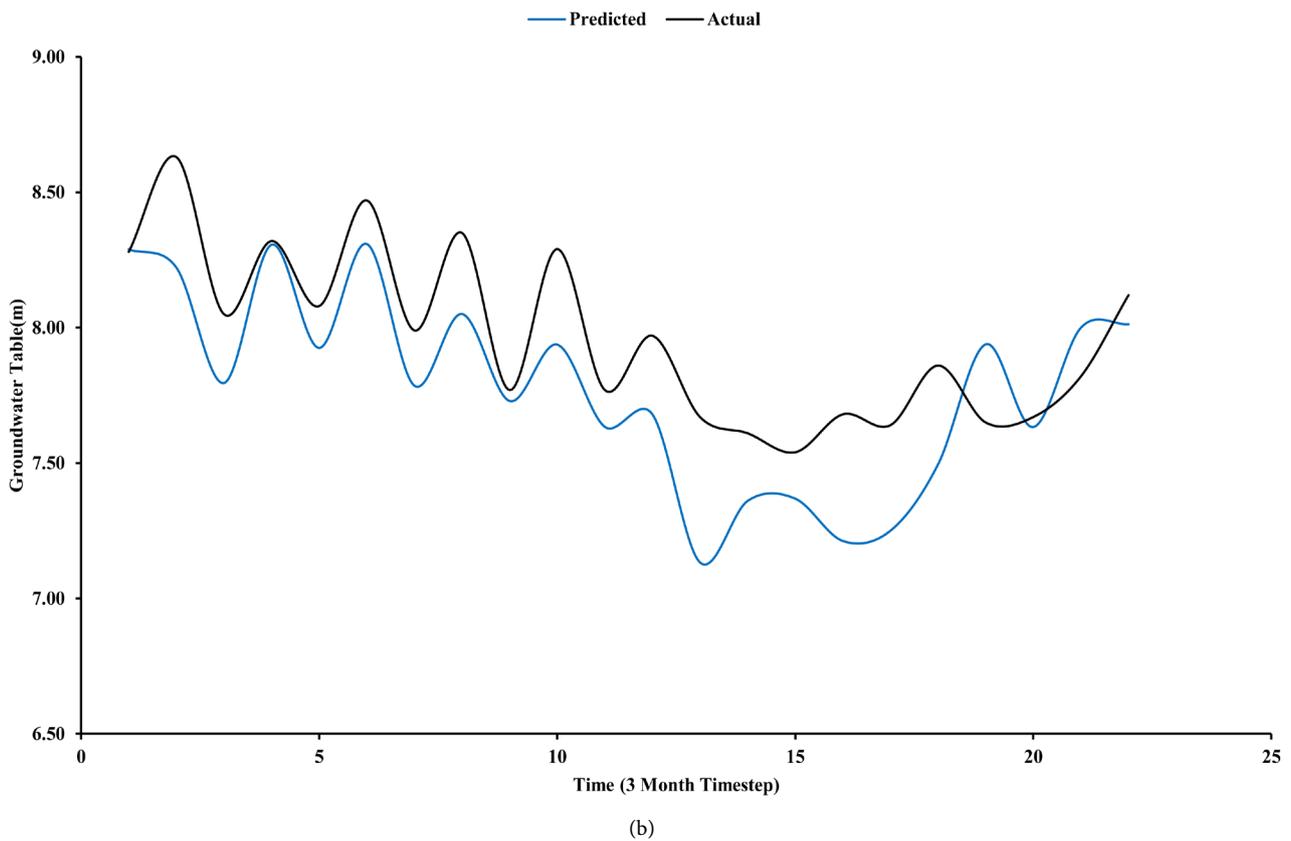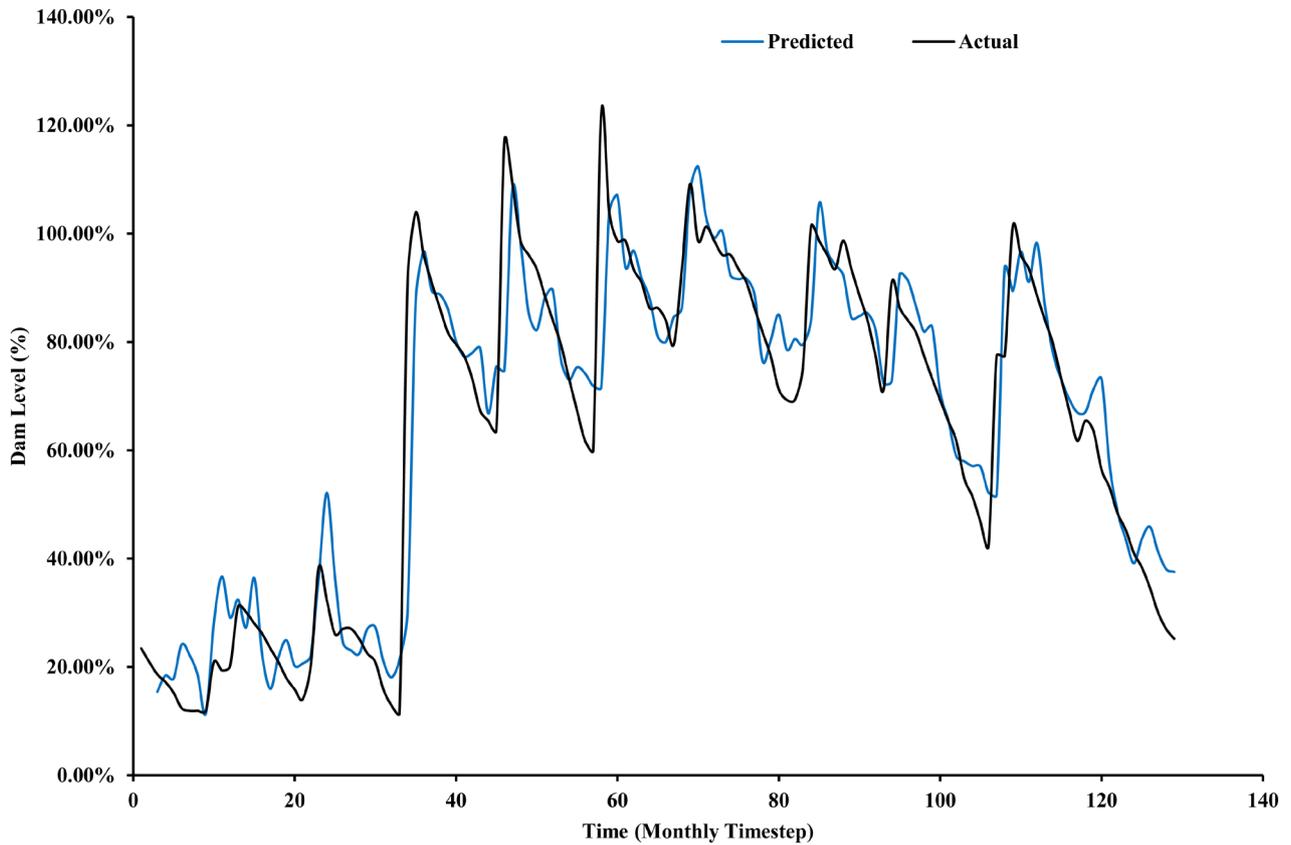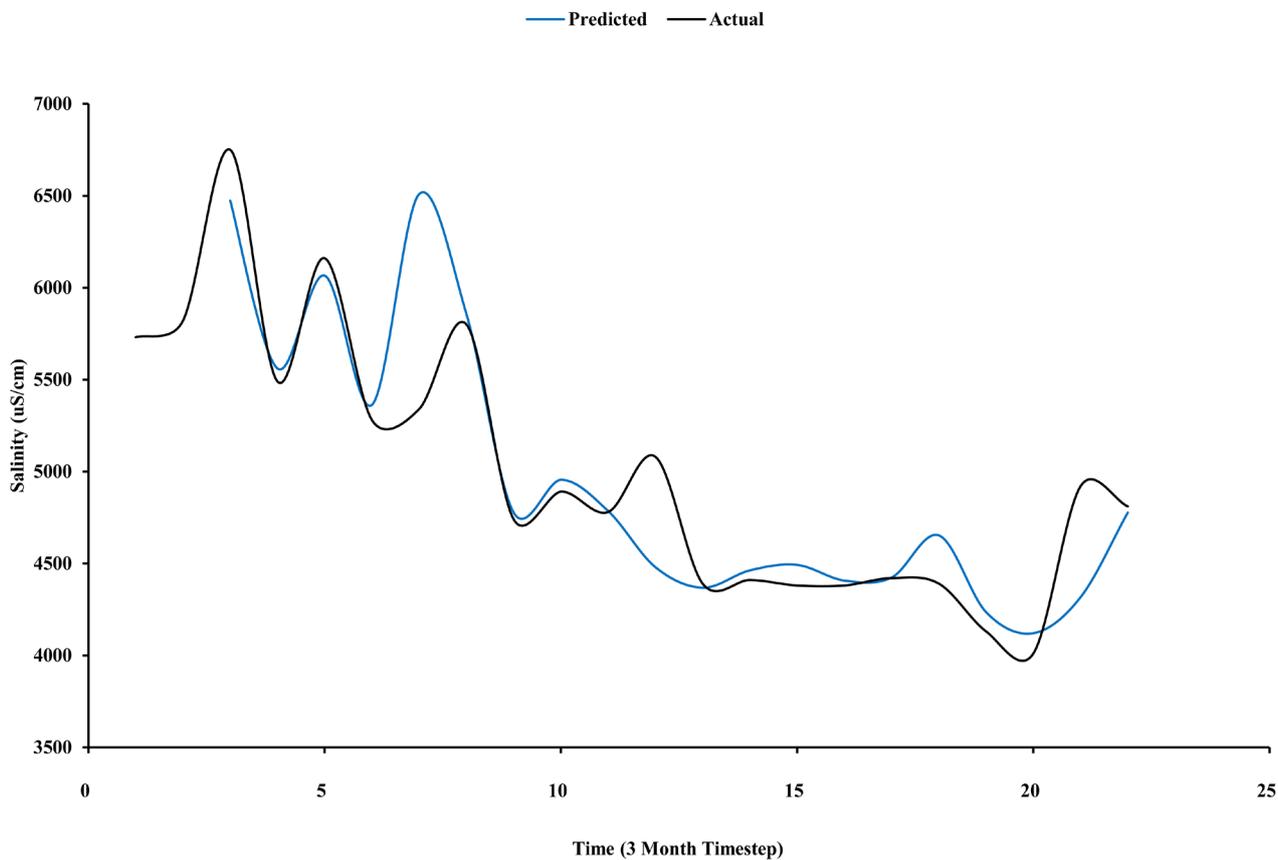
(c)
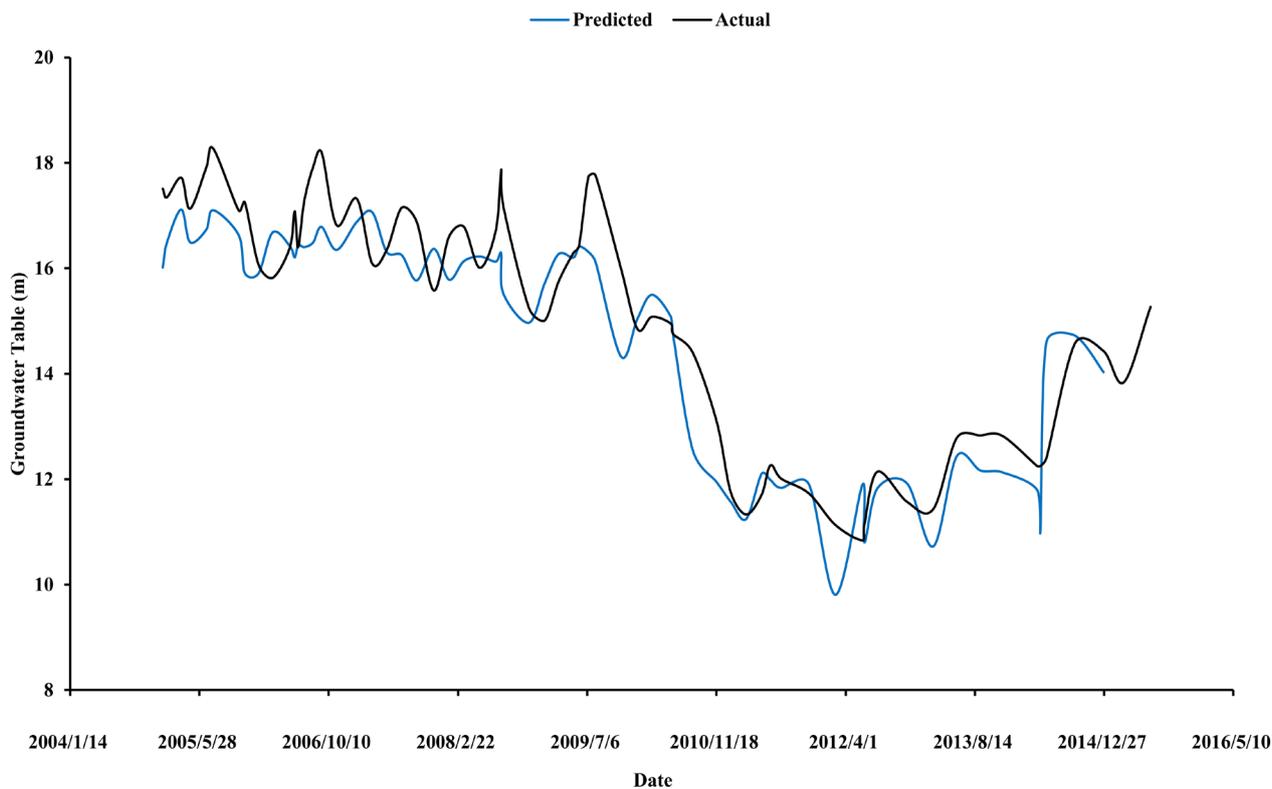


(d)

**Figure 9.** Prediction of hydrological drought; (a) prediction model 5; (b) prediction model 6; (c) prediction model 7; (d) prediction model 8; (e) prediction model 9.

The error occurring in this test verifies the simplicity of the relationship between groundwater levels and SPI values with an approximate accuracy of 97 percent. However, this model does not directly assess drought conditions, rather merely relates ground water fluctuations with respect to drought conditions. A greater assessment of water storage capability and depth capabilities would close this limitation. However, the results from the test may then be also used to predict salinity levels in the coastal groundwater aquifer.

### Model 7 (Salinity) Bore: 11700027B

The use of the predicted dataset from model 6 was utilized to assess the salinity changes, and the effects of water levels on these values. It is expected that the salinity levels will increase when ground water levels decrease. This becomes increasingly obvious when comparing Figure 9(b) with Figure 9(c) as the salinity levels tend to drop as ground water level decreases. This is due to saltwater intrusion from the interaction with the close proximity seawater. As the freshwater head decreases, salt water tends to seep in and start to replenish the reduced water level, which in turn, increases the salinity of the ground water. Therefore, to assess this, a FFB ANN model was utilized to find the relationship between the ground water level and the salinity content without needing to consider a time step.

The results obtained from the FFB show relatively sound performance and an accurate relationship. Despite the anomaly occurring at the 7th and 12th time step, the trend between the two series is more or less accurate. By using this tool

after predicting the ground water level, a quick assessment of the bore water quality can be approximately assessed 3 months in advance to ensure the salinity level does not exceed the (1000 uS/cm) mark. As can be seen in the graph, the salinity for this water far surpasses the minimum salinity requirement. This bore cannot be used effectively for watering purposes, however, was used to depict the relationship between groundwater levels and salinity. The model, however, can be increasingly accurate with the inclusion of more data points and therefore assisting the training function. Despite this, a clear relationship can be gathered with small amounts of data training. Average test error is 3.73 percent meaning that the resulting model is around 96% accurate.

### Model 8 (Ground Water: Groundwater Table) Bore: 11800007A

This prediction model assesses the groundwater table (depth to datum) of an inland aquifer in the Ross River basin. The results given in Figure 9(d) provide the representation of the bore depth with respect to a particular date. As previously discussed, many inputs were trialed however, the use of SPI proved to be the simplest and utmost effective to minimize error. As can be seen in Figure 9(d), the predicted dataset follows a direct correlation to the actual ground water levels. It is worth discussing the fact that this model had a substantially larger input dataset in comparison to that for the coastal aquifer. This enhanced the ability of the model to react accordingly with input deviations. It also allows the user to trust the predicted value because of the trained dataset. Despite this, minor discrepancies can be seen occurring within the dataset commonly at major points of flux.

The model gives an excellent 3-month ahead prediction with an accuracy of 95%. There are a couple of possible justifications behind the accuracy of this model including the underground contribution of stream flow from the Ross. This flow is highly dependent on precipitation and the seepage into the ground water streams. It can then be identified that the use of SPI as an input is an exceptional parameter. Also being in a suburban area, the number of bores and approximate water usage is relatively low when comparing to bores on agricultural estates. The usage affect is therefore minimal which, is another possible explanation for the accuracy of the prediction.

### Model 9 (Salinity) Bore: 11800007A

By using the predicted data set obtained from the prediction model 8, a FFB model was utilized to try to find a close relationship between salinity and groundwater levels. However, when viewing Figure 9(e), there seems to be no direct trend for the most part of the model. This model was carried out in the same manner, as model 7 however was unable to find direct correlations between results. One possibility may be due to the low amount of data points required for training. However, the previous model predicted well with a similar amount of data. The most probabilistic cause is because this bore does not sit close to seawater. Therefore, the salinity does not depend on the depth of the bore, as there is truly little or no salt water to intrution in this part of this aquifer. This however does not include contaminants and other intrusions on the aquifer, which

may very well depend on the height on the bore. Furthermore, the salinity level of this bore is sufficiently low and can be used for most of the vegetation.
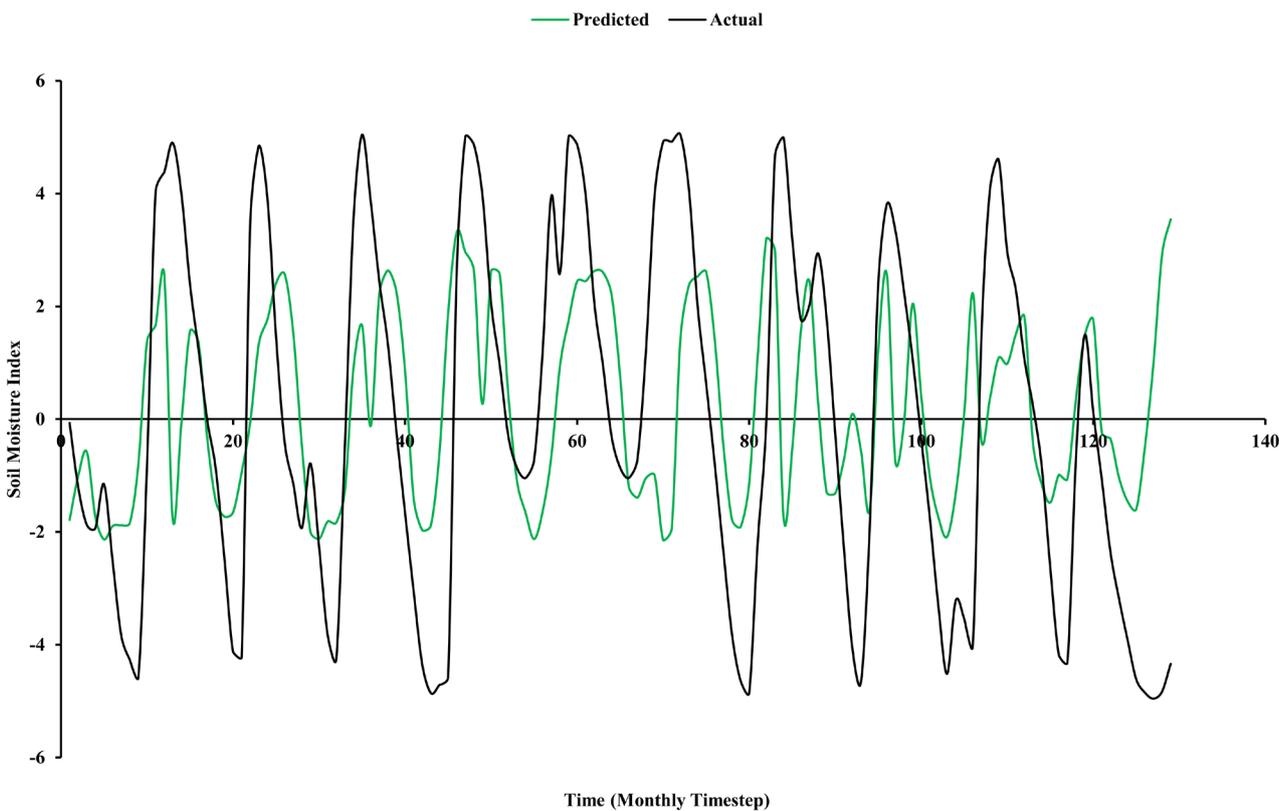
## 4.3. Agricultural Drought

For the assessment of agricultural drought, two models were utilized which includes both meteorological and agricultural input datasets. To help reduce error and simplify these models, a conversion of root-zone soil moisture to the drought index SMI was conducted.

The first prediction model generated took advantage of the meteorological index, SPI. The use of the SPI simplifies the models input requirements and alerts the model of the wetter and drier periods. The results obtained from the simulation (model 10), as seen in Figure 10(a), it does show that the SPI gives a good indication on the wetter and drier periods. However, the amplitude of the fluctuations in the predicted dataset is lower than that of the actual dataset. This occurs particularly where the SMI is greater than 2.0 in the actual dataset. The reasoning for this may well be that because, this dataset is of root-zone soil moisture therefore when precipitation occurs, sufficient seepage is needed to be able to reach this layer of soil. Since the seepage rate is slow, possible evaporation may reduce the amount of water in the top soil moisture therefore reducing the amount of moisture to seep to the root zone. However, in reverse effect of this in the dry periods, the intensity can be mimicked. Because of this, the training program may have found that minimal error occurred when focusing on the bottom end of data. The usual process to avoid this is to increase the number of neurons in the program. However, since there was only one input, minimal effect occurred and correlation between the two values stayed the same. Despite the intensity, there is a clear trend in the data, which can essentially show the approximate fluctuations of the soil moisture. Furthermore, the model follows an accurate time step arrangement where the peaks occurring in the predicted dataset matches the time step of the actual data. Although accurate readings are not absolute, the change in soil moisture can be assessed to some extent. With that in mind, a 3-months prediction can allow the agricultural industry to prepare and control soil conditions to the appropriate and desired growing capabilities. Regardless of the percentage error being 78 percent as seen in Figure 10(a), given the trend of this model, it does give a good indication on the soil moisture trend.

The second prediction model in this category (model 11) was developed based on the results obtained from the previous prediction model in the hope that the use of evapotranspiration will help reduce the error focus from the lower regions of the SMI. By observing Figure 10(b), there is a clear trend and accurate time step formation between the simulated dataset and the actual dataset. However, the simulation once again does not meet the intensity that occurs in the actual data set. Although this occurs, the hypothesis mentioned in in the preceding paragraph about the use of evapotranspiration to reduce the error focus on the lower SMI values has been validated. It does however reduce the ability to
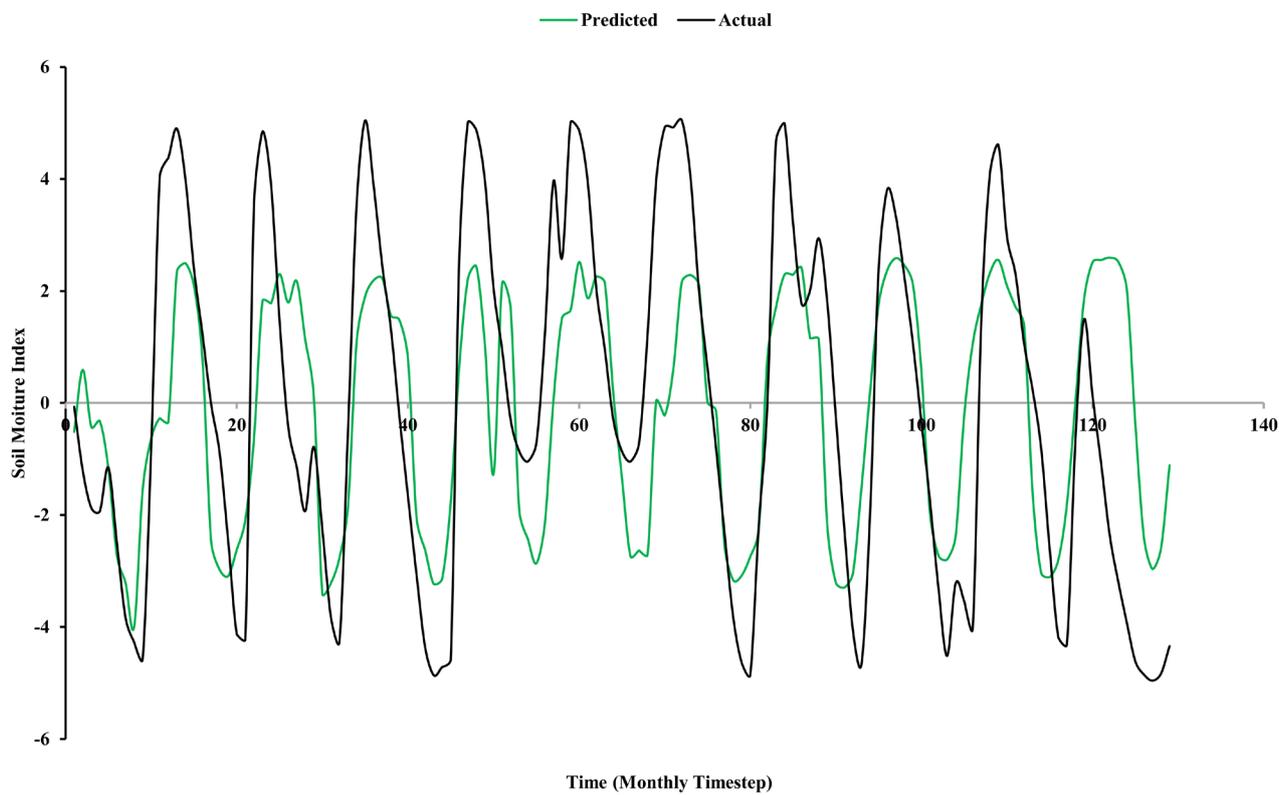
(a)



(b)

**Figure 10.** Agricultural drought prediction; (a) prediction model 10; (b) prediction model 11.

predict the scenarios because of the intensity change but does give a more even distribution when concerning the trend. At the final time steps in the simulation however, there appears to be the presence on an anomaly set of predictions. This simulation anomaly may be due to uneven precipitation occurrence in those months. The values for evapotranspiration were taken relatively close to the Ross River dam whereas; precipitation values were taken at the airport some 15 km inland. In these particular time steps, there was substantial rain occurring over the coast of Townsville, which did not occur over the dam. This was an anomalous rainfall occurrence and is the highly probable reasoning for these results. Despite this, there appears to be a strong trend in the model and the generation of the SMI can be achieved at a 3-month prediction efficiently.

## 5. Conclusions Limitations and Recommendations

Artificial Neural Network based prediction models can be utilized as an effective prediction tool for the prediction of possible drought scenarios. The developed models produced sound results that could essentially help the North Queensland economy deal with the effect of drought with the assistance of future prediction models. Based on the results of the present study, it can be concluded that ANN based modelling is an effective, simple, and efficient prediction tool to accurately forecast drought conditions. It is also demonstrated that meteorological drought conditions can effectively be assessed using the SPI which can be approximately forecasted for a lead time of up to 6 months in advance. Hydrological conditions have the ability to be forecasted using drought indices such as SMI and SPI. Surface water levels of the Ross River dam can be predicted at a lead-time of 3 months without any considerable error. Ground water conditions including groundwater levels and salinity intrusion for the specified bore locations of the coastal aquifer can also be successfully modelled using the proposed ANN based models. On coastal aquifers, salinity levels are highly dependent on ground water levels. However, groundwater level has minimal effect on saltwater intrusion for inland aquifers. SMI is an effective tool to assess agricultural drought conditions and can be forecasted using both meteorological and agricultural values with a lead time of 3 months. This can allow the agricultural industry to prepare and control soil conditions to the appropriate and desired growing capabilities.

For accurate prediction of drought scenarios using any prediction model requires a considerable amount of past data. Furthermore, the calculation of both SPI and SMI should utilize 20 - 25 years' worth of normalized data to ensure greater variations and values of the index. However, only 11 years of historical data were accessible for calculating the indices.

The management of the use of bores, particularly in gardens and lawns, heavily relies on the conservation programs that are in place. A good example of a conservation program is Townsville Water Program where the restriction of watering has set periods, *i.e.*, between 6 pm and 8 pm twice a week, which moreover depends on the restriction level. If these water restrictions are not abided by,

then fines may be dealt with. This however is not the most effective solution for groundwater control and needs to revise as the depreciation of water in aquifers often causes saltwater intrusion. Water quality in aquifers can be influenced by the effect of saltwater intrusion as the water level of the aquifer falls. Appropriate management of bores needs to be implemented to restrict this event from occurring. Two common alternatives for controlling saltwater intrusions include a planned pumping strategy, and planned extractions from barrier wells [21].

Salinity has been recognized as a problem around coastal regions of Australia in aquifers. Salinity affects crop and vegetation growth when used for irrigation purposes and in high enough doses, can reverse effect osmosis. Acceptable levels of salinity are usually of salinity levels of 700 μS/cm or less. However, the effect it has on individual crop yield depends on the crop itself. For example, fruit crops can accept salinity of up to 1000 μS/cm without affecting crop yield and pastures can accept up to 1300 μS/cm without affecting yield. All of these aspects will need to be considered for developing short term groundwater management strategies.

The main focus of this study was to construct and then verify the feasibility and plausible accuracy of predicting various drought scenarios, including Meteorological, Hydrological, and Agricultural drought scenarios on a sub-regional scale. A number of different prediction models are developed and the performance of these models is evaluated, within the limits of available data, related to the tropical application area of Townsville in Queensland, in Northern Australia. The choice of the study area was based on the recent occurrence of a fairly long drought (2014-2015 to early 2019) that had serious consequences in this study area comprising of urban and agricultural land use. The prediction models developed and evaluated predict different likely consequences of drought scenarios related to different aspects. The variations tested represent advantages and disadvantages, while affecting accuracy. Therefore, no particular model is selected or recommended for universal use.

The aim was to analyse the feasibility of using the proposed approach for prediction of different drought aspects. Further studies with robust data sets are necessary for recommending a particular set of models with wide applicability. Also, due to the fact that the application is limited to a region forming part of a state, it was felt that the utility of incorporating larger scale models, e.g., Global Circulation Models (GCMs) may not be very useful for this limited application area, short-term prediction scenario. The lead times were however, short enough to be useful as real time prediction models, with possible utilization for short term or real time planning and operation of reservoir (e.g., Ross River dam), or, for developing real-time groundwater management strategies to mitigate impending drought consequences.

This study is a preliminary effort in developing drought prediction models for the dry tropical region in the vicinity of Townsville. More rigorous development and evaluation of prediction accuracy are no doubt necessary as the next step in order to incorporate these prediction models for evolving sustainable, reliable

and efficient regional scale groundwater and surface water management strategies under extreme or, catastrophic conditions. Capability of predicting impending droughts with reasonable lead time and accuracy can also help develop optimal water management strategies based on planned releases from reservoirs.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

[1]  Mishra, A.K. and Singh, V.P. (2010) A Review of Drought Concepts. *Journal of Hydrology*, **39**, 202-216. https://doi.org/10.1016/j.jhydrol.2010.07.012

[2]  Hao, Z. and Singh, V.P. (2015) Drought Characterization from a Multivariate Perspective: A Review. *Journal of Hydrology*, **527**, 668-678. https://doi.org/10.1016/j.jhydrol.2015.05.031

[3]  McCabe, G.J., Betancourt, J.L., Gray, S.T., Palecki, M.A. and Hidalgo, H.G. (2008) Associations of Multi-Decadal Sea-Surface Temperature Variability with Us Drought. *Quaternary International*, **188**, 31-40. https://doi.org/10.1016/j.quaint.2007.07.001

[4]  Dai, A. (2011) Drought under Global Warming: A Review. *Wiley Interdisciplinary Reviews: Climate Change*, **2**, 45-65. https://doi.org/10.1002/wcc.81

[5]  Pedro-Monzonís, M., Solera, A., Ferrer, J., Estrela, T. and Paredes-Arquiola, J. (2015) A Review of Water Scarcity and Drought Indexes in Water Resources Planning and Management. *Journal of Hydrology*, **527**, 482-493. https://doi.org/10.1016/j.jhydrol.2015.05.003

[6]  Kumar, V. and Panu, U. (1997) Predictive Assessment of Deverity of Agricultural Droughts Based on Agro-Climatic Factors. *JAWRA Journal of the American Water Resources Association*, **33**, 1255-1264. https://doi.org/10.1111/j.1752-1688.1997.tb03550.x

[7]  Lee, Y.-S. and Tong, L.-I. (2011) Forecasting Time Series Using a Methodology Based on Autoregressive Integrated Moving Average and Genetic Programming. *Knowledge-Based Systems*, **24**, 66-72. https://doi.org/10.1016/j.knosys.2010.07.006

[8]  Liu, W.T. and Juárez, R.I.N. (2001) Enso Drought Onset Prediction in Northeast Brazil Using NDVI. *International Journal of Remote Sensing*, **22**, 3483-3501. https://doi.org/10.1080/01431160010006430

[9]  Mishra, A.K. and Desai, V.R. (2005) Drought Forecasting Using Stochastic Models. *Stochastic Environmental Research and Risk Assessment*, **19**, 326-339. https://doi.org/10.1007/s00477-005-0238-4

[10]  Mishra, A.K. and Singh, V.P. (2011) Drought Modeling—A Review. *Journal of Hydrology*, **403**, 157-175. https://doi.org/10.1016/j.jhydrol.2011.03.049

[11]  Rao, A.R. and Padmanabhan, G. (1984) Analysis and Modeling of Palmer's Drought Index Series. *Journal of Hydrology*, **68**, 211-229. https://doi.org/10.1016/0022-1694(84)90212-9

[12]  Ferna'ndez, C., Vega, J.A., Fonturbel, T. and Jime'nez, E. (2009) Streamflow Drought Time Series Forecasting: A Case Study in a Small Watershed in North West Spain. *Stochastic Environmental Research and Risk Assessment*, **23**, Article No. 1063. https://doi.org/10.1007/s00477-008-0277-8

[13]  Kuligowski, R.J. and Barros, A.P. (1998) Experiments in Short-Term Precipitation

Forecasting Using Artificial Neural Networks. *Monthly Weather Review*, **126**, 470-482. https://doi.org/10.1175/1520-0493(1998)126<0470:EISTPF>2.0.CO;2

[14] Morid, S., Smakhtin, V. and Bagherzadeh, K. (2007) Drought Forecasting Using Artificial Neural Networks and Time Series of Drought Indices. *International Journal of Climatology*, **27**, 2103-2111. https://doi.org/10.1002/joc.1498

[15] Steinemann, A. (2003) Drought Indicators and Triggers: A Stochastic Approach to Evaluation. *JAWRA Journal of the American Water Resources Association*, **39**, 1217-1233. https://doi.org/10.1111/j.1752-1688.2003.tb03704.x

[16] Guttman, N.B. (1999) Accepting the Standardized Precipitation Index: A Calculation Algorithm. *JAWRA Journal of the American Water Resources Association*, **35**, 311-322. https://doi.org/10.1111/j.1752-1688.1999.tb03592.x

[17] Guttman, N.B. (1998) Comparing the Palmer Drought Index and the Standardized Precipitation Index. *JAWRA Journal of the American Water Resources Association*, **34**, 113-121. https://doi.org/10.1111/j.1752-1688.1998.tb05964.x

[18] Hunt, E.D., Hubbard, K.G., Wilhite, D.A., Arkebauer, T.J. and Dutcher, A.L. (2009) The Development and Evaluation of a Soil Moisture Index. *International Journal of Climatology*, **29**, 747-759. https://doi.org/10.1002/joc.1749

[19] Dai, A., Trenberth, K.E. and Qian, T. (2004) A Global Dataset of Palmer Drought Severity Index for 1870-2002: Relationship with Soil Moisture and Effects of Surface Warming. *Journal of Hydrogeometry*, **5**, 1117-1130. https://doi.org/10.1175/JHM-386.1

[20] Latif, M., Barnett, T.P., Cane, M.A., Flügel, M., Graham, N.E., von Storch, H., Xu, J.-S. and Zebiak, S.E. (1994) A Review of ENSO Prediction Studies. *Climate Dynamics*, **9,** 167-179. https://doi.org/10.1007/BF00208250

[21] Dhar, A. and Datta, B. (2009) Saltwater Intrusion Management of Coastal Aquifers. I: Linked Simulation-Optimization. *Journal of Hydrologic Engineering*, **14**, 1263-1272. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000097