Article

Predicting the Effects of Climate Change on Water Temperatures of Roode Elsberg Dam Using Nonparametric Machine Learning Models

Thalosang Tshireletso * , **Pilate Moyo and Matongo Kabani**

Department of Civil Engineering, University of Cape Town, Cape Town 7700, South Africa; pilate.moyo@uct.ac.za (P.M.); matongo@uct.ac.za (M.K.)
* Correspondence: TSHTHA070@myuct.ac.za

Abstract: A nonparametric machine learning model was used to study the behaviour of the variables of a concrete arch dam: Roode Elsberg dam. The variables used were ambient temperature, water temperatures, and water level. Water temperature was measured using twelve thermometers; six thermometers were on each flank of the dam. The thermometers were placed in pairs on different levels: avg6 (avg6-R and avg6-L) and avg5 (avg5-R and avg5-L) were on level 47.43 m, avg4 (avg4-R and avg4-L) and avg3 (avg3-R and avg3-L) were on level 43.62 m, and avg2 (avg2-R and avg2-L) and avg1 (avg1-R and avg1-L) were on level 26.23 m. Four neural networks and four random forests were cross-validated to determine their best-performing hyperparameters with the water temperature data. Quantile random forest was the best performer at mtry 7 (Number of variables randomly sampled as candidates at each split) and RMSE (Root mean square error) of 0.0015, therefore it was used for making predictions. The predictions were made using two cases of water level: recorded water level and full dam steady-state at Representative Concentration Pathway (RCP) 4.5 (hot and cold model) and RCP 8.5 (hot and cold model). Ambient temperature increased on average by 1.6 °C for the period 2012–2053 when using recorded water level; this led to increases in water temperature of 0.9 °C, 0.8 °C, and 0.4 °C for avg6-R, avg3-R, and avg1-R, respectively, for the period 2012–2053. The same average temperature increase led to average increases of 0.7 °C for avg6-R, 0.6 °C for avg3-R, and 0.3 °C for avg1-R for a full dam steady-state for the period 2012–2053.

Keywords: machine learning; climate change; dam safety; water temperature

1. Introduction

Dams are an important infrastructure, and their failure has high economic and social consequences. This is because they have an associated risk that must be managed in a continuous and updated process [1]. In the context of dam safety, risk is estimated by combining the impact of a scenario, the probability of occurrence of that scenario, and the associated consequence of that scenario [2]. Dam safety is about managing dams to a safety level that minimizes any risk to life, property, essential services, and the environment [3]. Normally, the associated risk posed on dams by a scenario is carried out assuming stationary climatic and non-climatic conditions. However, the projected alterations due to climate change are likely to affect different factors driving dam risk [4], like variations in extreme temperatures or frequency of heavy precipitation events.

Risk assessment is the methodology that is used to manage dam risk; it is a useful methodology that incorporates traditional and state-of the-art methods to accomplish dam safety in an accountable and comprehensive way [5]. The development and application of risk-assessment techniques worldwide in the dam industry [3] has abetted safety governance and supports decision making in the adoption of structural and non-structural risk reduction measures. There are different ways of managing the risk, namely the following:

- The traditional approach- relies on experience, technical guidelines, and engineering judgement [6].
- The standards-based approach- relies on safety factors to assess the level of stability, stress, durability, and deformation of parts of the dam and their foundation [7].
- The risk-based approach- is a systematic and scientific evaluation of the likelihood of occurrence of an adverse event that may compromise the safety of a dam and the associated consequences of dam failure.

Traditionally, dam safety has been achieved using the traditional approach, which relies on experience. The standards approach has also been used; it requires dams to withstand certain defined loads. The traditional approach plays a key role in the design of new dams as well as assessment of existing dams in Vietnam [7]; however, it has limitations in terms of indirectly assessing dam failure as well addressing input uncertainties. This approach assumes a stationary condition in the variability in climate change [8], including the frequency and magnitude of extreme events [9]. Changes in climate factors such as extreme temperature, frequency of heavy precipitation events [10], and increasing average temperatures are likely to affect the different factors driving dam [7] by changing the loads acting on dams to what dams were not designed for, proving that the assumptions of a stationary baseline are no longer appropriate for long-term dam safety management.

The global effect of climate change on dam risk must be assessed through the combination of various projected effects influencing each dam-safety aspect, considering their interdependencies rather than by a simple accumulation of separate impacts [1]. Risk-based dam safety can help address such analyses; the approach is accountable and comprehensive [7] because it combines three aspects: what can happen (infrastructure failure), how likely it is to happen (failure probability), and what are its consequences (failure consequences) [11].

Reference institutions have developed guidelines that account for climate change in their decision-making strategies, but climate change information is vast and scattered and its application to specific analyses such as dam-safety assessments remains a challenge [1]. Most studies focus only on how climate change will impact hydrological loads [4], ignoring other loads such as ambient temperature and water temperature. Other studies with a wider scope only reach qualitative assessments [12].

Temperature has direct effects on thermoelastic properties of concrete and can affect creep and alkali aggregate reaction [13]. Arch dams experience thermal heat through ambient temperature via radiation and water temperature. Ambient temperature and its effects on behaviour have been widely studied, while the effects of water temperature have received less attention from researchers.

Figure 1 shows the plot of the daily time-series of ambient temperatures, water temperatures at different heights, and the water level of Roode Elsberg dam. The effects of ambient temperature are felt when the water level is low; this effect is coupled by water temperatures below the water level. Water temperature drives the behaviour of the dam mostly when the dam is full.
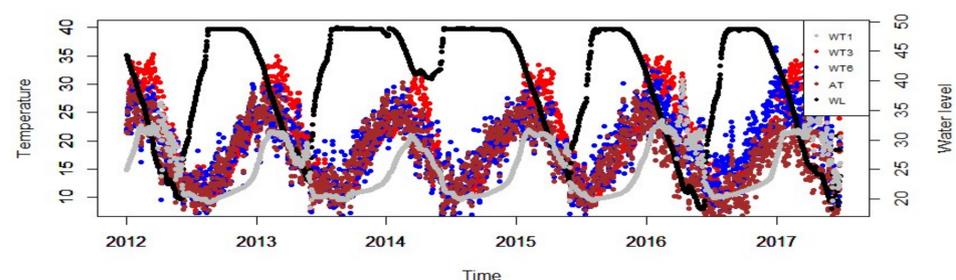
Figure 1. Roode Elsberg dam monitoring data 2012–2017 (water temperature 1 (WT1), water temperature 3 (WT3), water temperature 6 (WT6), ambient temperature (AT), and water level (WL)), temperature in °C.

Figure 2 shows the plot of a daily time-series of ambient temperatures, water temperatures at different heights, and radial deformations of Roode Elsberg dam.
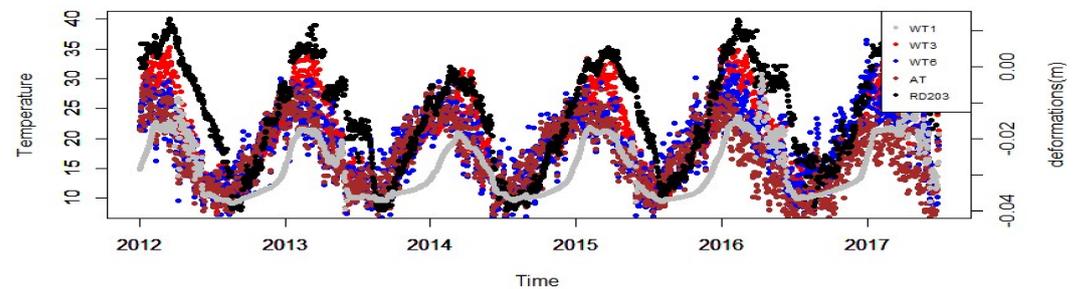
Figure 2. Roode Elsberg dam monitoring data 2012–2017 (water temperature 1 (WT1), water temperature 3 (WT3), water temperature 6 (WT6), ambient temperature (AT), and radial deformations 203 (RD203)), temperature in °C.

From Figures 1 and 2, it can be concluded that water level drives the downstream movement of the dam whilst temperature drives the upstream movement of the dam. When the dam fills up, it moves downstream. The water body is then heated by ambient temperature, which makes the water temperature increase. This makes the dam expand, hence moving upstream. The effects of water temperature have been shown to be more important than the effects of ambient temperature in temperature distribution of the dam across the thickness [13]. In order to understand the risk that climate change poses on dam safety, water temperatures must be incorporated in risk assessment as well as in model predictions.

The aim of this work is to understand how different climate-change permutations will influence water temperatures of Roode Elberg by mid-century (2053). This work will concentrate on how different changes in ambient temperatures affect the water temperature of the water body by 2053. This was archived by manipulating the relationship between ambient temperature and water temperature of Roode Elsberg dam using machine-learning data models. The periods of study were chosen to be 2012–2017 and 2048–2053. Predictions were made at Representative Concentration Pathway 4.5 (RCP 4.5) and at Representation Concentration Pathway 8.5 (RCP 8.5).

2. Materials and Methods

The study was based on Roode Elsberg dam, located in Western Cape, South Africa. The monitoring system included measurement of environmental and operational conditions, i.e., water temperatures, measured by thermometers embedded on the dam wall; water level; deformations; and accelerations. The monitoring was carried out by CoMSIRU/UCT and the Department of Water and Sanitation.

2.1. Roode Elsberg Dam

Roode Elsberg dam is in the Western Cape province, South Africa, about 130 km northeast of Cape Town near the town of de Doorns, at coordinates (33.4361° S, 19.5680° E), Figure 3. Construction of the dam was completed in 1969, and its main purpose is irrigation of vineyards in the surrounding farms and limited domestic use via a 7-km tunnel. The dam is a double curvature concrete arch dam with a centrally located spillway and gross capacity of 8.21 million m³. The height to the lowest foundation point is 72 m, and the length of the crest is 274 m, consisting of two galleries, one following the foundation level and one top instrumentation gallery located about 20 m above the foundation level.

Figure 3. Roode Elsberg dam.

2.1.1. Roode Elsberg Dam-Monitoring System

In order to understand the behaviour of Roode Elsberg, two monitoring systems were installed on the dam. These included a continuously monitored GPS system at four survey beacons in 2010 and a dynamic monitoring system at the dam crest in 2013. In addition, environmental and operational conditions were measured, i.e., water level measured using staff gauges and water temperatures; a weather station; and a suite of thermometers located at different water levels: 26.23 m, 46.62 m, and 47.30 m. Figure 4 shows the layout of the Roode Elsberg GPS monitoring systems, where blue indicates control stations P01 and P02 while red indicates rover stations P203 and P206 on the left and right flanks, respectively.

Figure 4. Roode Elsberg GPS monitoring systems.

2.1.2. Measurement of Environmental and Operational Factors

The environmental and operational factors include ambient temperature, water temperature, and water level. Ambient temperature is measured by the weather station installed on the dam crest, and water levels are measured by staff gauges. There are also thermometers that are embedded into the dam wall at different levels to measure water temperatures; Figure 5 below shows the position of water temperatures on the dam wall. There are 6 thermometers on each side of the wall; avg1-R indicates that thermometer 1 is on the right flank and avg1-L indicates that thermometer 1 is on the left flank. Avg1-R and avg2-R are on the same level, avg3-R and avg4-R are on the same level, and avg5-R and avg6-R are on the same level, as indicated by Figure 5. This applies to thermometers on the left flank.
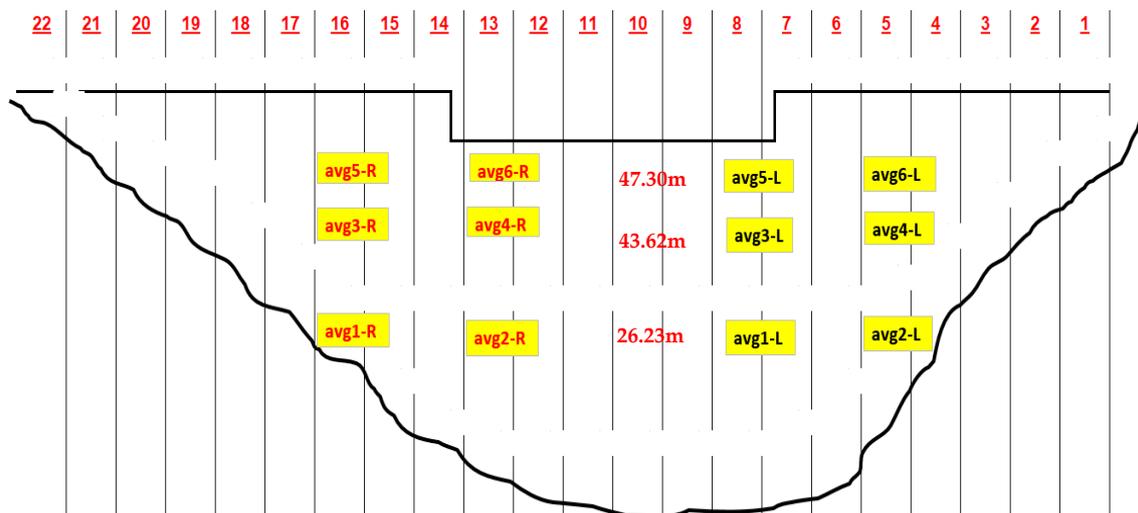
Figure 5. Wall thermometers measuring water temperature.

2.2. Roode Elsberg Monitoring Data

The monitoring data used for this document were the water temperatures (2011–2020), water level (1969–2017), and ambient temperatures (1979–2020). The water temperature and water level were measured on-site, while the ambient temperature was downloaded from http://climexp.knmi.nl/selectdailyseries.cgi?id=someone@somewhere for a town in Western Cape called de Doorns, which is less than 5 km from the dam site. The ambient temperature was downloaded because the weather station data was not available; the closest town was de Doorns, which happens to be in a valley just like Roode Elberg dam.

Figures 6–11 show plots of water temperatures for the left against the right flanks. Thermometers avg1 (avg1-R and avg1-L) and avg2 (avg2-R and avg2-L) for both flanks exhibited similar temperature patterns (Figures 6 and 7). This similarity is shown by temperatures recorded by avg3 and av4 for both flanks, which are on the same level, as shown by Figures 8 and 9. For thermometers avg5 and avg6, there was a slight difference in the temperatures recorded. Avg5-L showed higher values than avg5-R in summer (November–January; Figure 10), whilst avg6-L showed slightly lower values compared to avg6-R in winter (June–July; Figure 11).
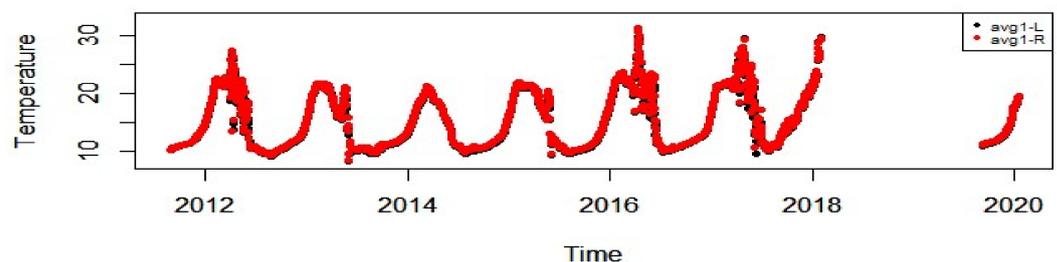

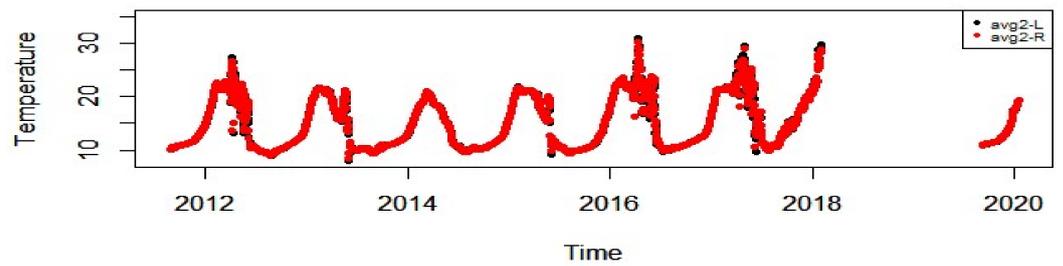**Figure 6.** Avg1-R against avg1-L (Temperature in °C).

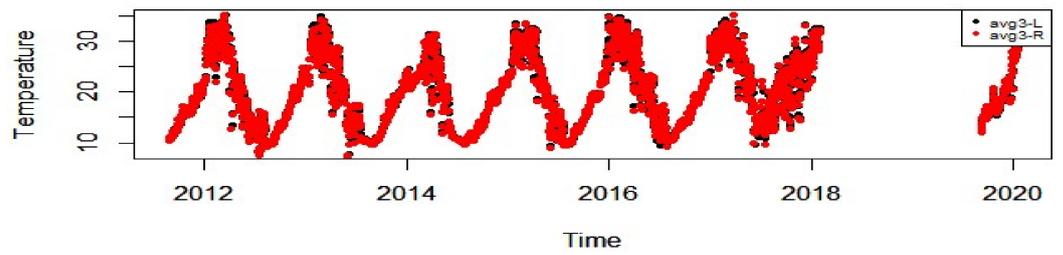**Figure 7.** Avg2-R against avg2-L (Temperature in °C).

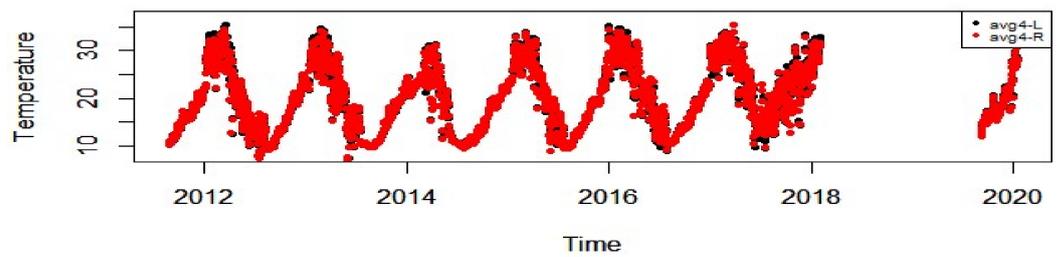**Figure 8.** Avg3-R against avg3-L (Temperature in °C).

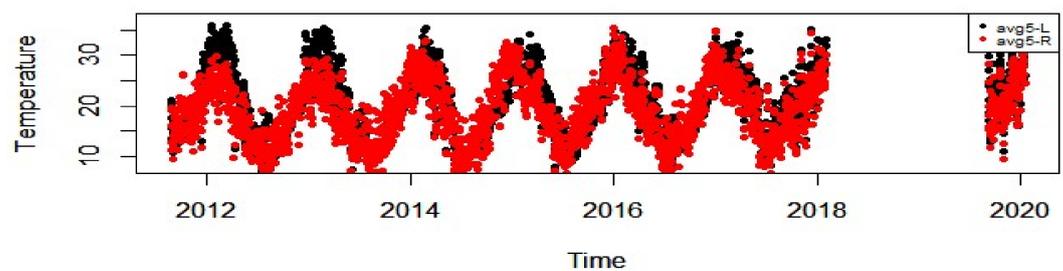**Figure 9.** Avg4-R against avg4-L (Temperature in °C).

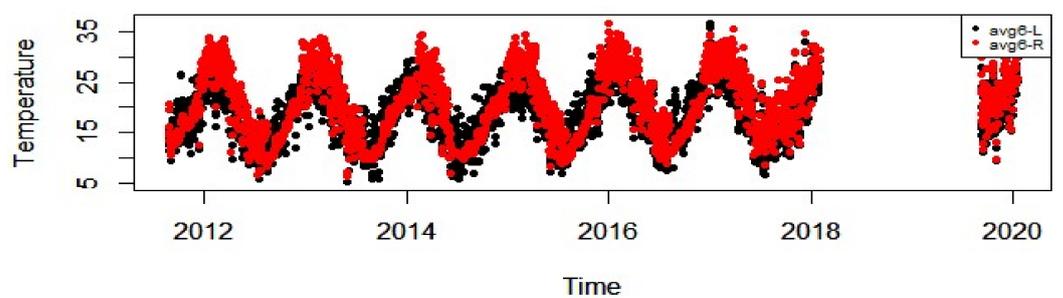**Figure 10.** Avg5-R against avg5-L (Temperature in °C).

Figure 11. Avg6-R against avg6-L (Temperature in °C).

The slight differences in the water temperature time-series shown by Figures 10 and 11 were caused by the location of the dam relative to the sun's trajectory. Roode Elsberg is

in a valley. The sun's trajectory is on the upstream side of the dam wall, with mountains on either side of the valley. Thermometers avg5 and avg6 were reached by the sun at different times of the dam. This behaviour was not recorded by the other thermometers, avg1–avg4, because of their locations on the dam wall.

For the purpose of this study, data from the right flank will be used. Since avg1-R and avg2-R, avg3-R and avg4-R, and avg5-R and avg6-R exhibit similar patterns, avg1-R, avg3R, and avg6-R will be used for analysis.

Figure 12 shows the ambient temperatures for de Doorns, which is less than 5 km from the dam site. The water level was recorded at Roode Elsberg dam. The calculated decadal change in temperature from 1979–2020 shows a temperature increase of 0.7 °C. The water level was not recorded daily until the year 2012. Figure 13 shows the daily water level for Roode Elsberg dam from 2012–2017.
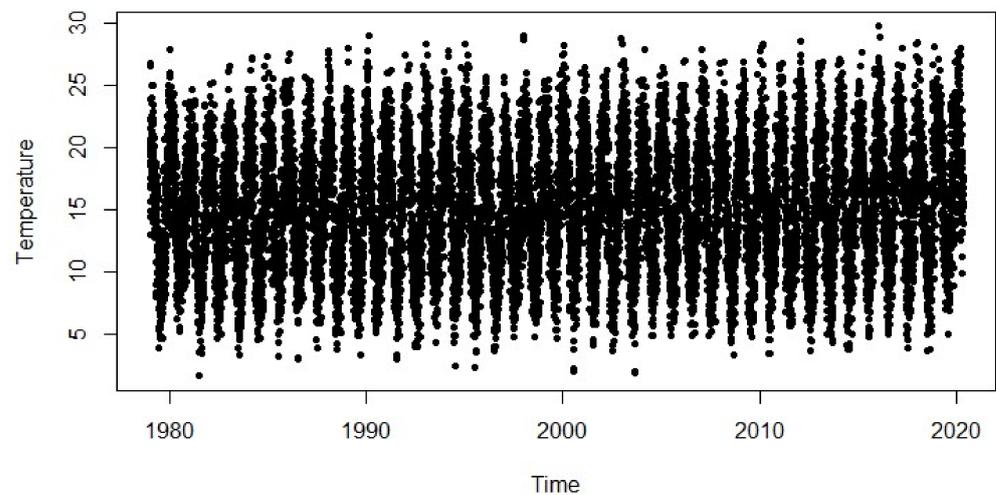

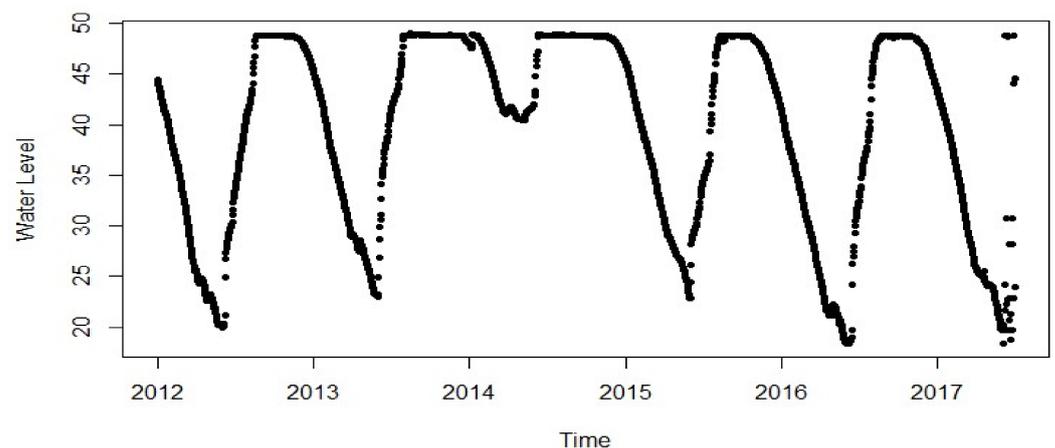**Figure 12.** de Doorns ambient temperature 1979–2020 (Temperature in °C).

Figure 13. Roode Elberg water level 2012–2017 (water level in metres(m)).

Variable Relationships

In order to be able to build a robust machine-learning model to predict water temperatures, the relationships of the variables were studied. The following figures show the relationships studied.

Figure 14 shows the relationship between Roode Elsberg's dam water temperature for 60-day moving averages and water levels. A 60-day moving average is used because it takes on average 53 days for the dam to respond to temperature increases. The moving average increases when the water level is low and decreases when the dam fills up.
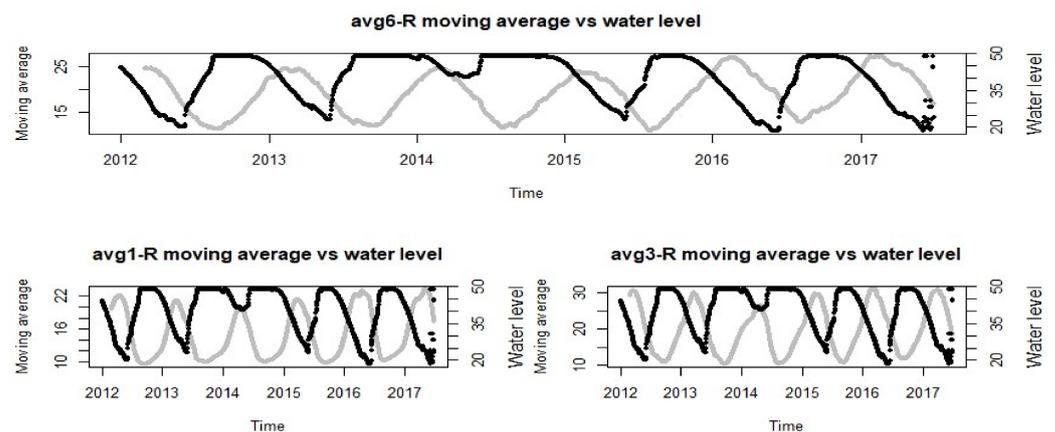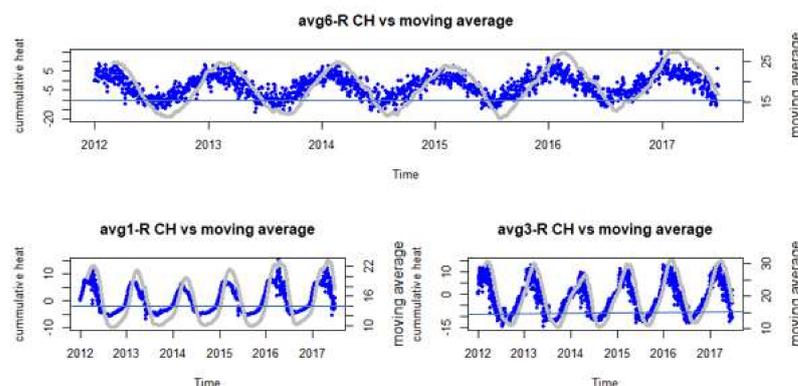
Figure 14. Avg6-R, avg1-R, and avg3-R 60-day moving averages (°C) (grey) against water level (m) (black).

Figure 15 shows the moving averages for avg6-R, avg1-R, and avg3-R. The horizontal blue line shows the closure temperature of the dam, which is approximately 15.5 °C. When temperatures are above this line, the dam is expanding, and when they are below the line, the dam is contracting. The vertical lines show the peaks of the moving averages. In most cases, avg1-R and avg3-R peak almost at the same time, except in 2014, when the dam did not lose a lot of water. This shows that the influence of water temperatures at different levels of the dam varies with time.
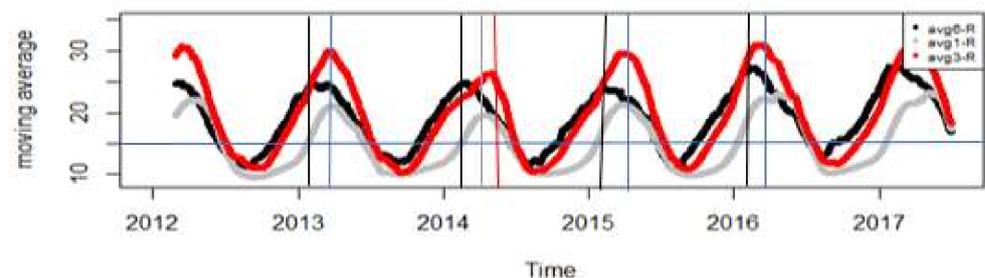
Figure 15. Avg5-R, avg1-R, and avg3-R 60-day moving averages (°C).

Figure 16 shows cumulative heat against the 60-day moving average. The cumulative heat is calculated with respect to 1 January 2012. The blue horizonal lines show the closure temperature of the dam, which is approximately 15.5 °C. The figure shows that, for most of the year, the dam is heated (expanding) more than cooled (contracting); this varies at different depths of the dam. Avg6-R, avg1-R, and avg3-R have amplitudes on average of 7.5 °C, 10 °C, and 15 °C, respectively, above the closure temperature.

Figure 16. Avg6-R, avg1-R, and avg3-R cumulative heat (°C) (blue) against the 60-day moving average (°C) (grey).

Figure 17 shows the 60-day moving variance against the water level. There is no relationship between the avg6-R moving variance and the water level. This is supported by the fact that the thermometer is almost at the dam crest. There is a relationship between the 60-day moving variances of avg1-R and avg3-R and the water level.
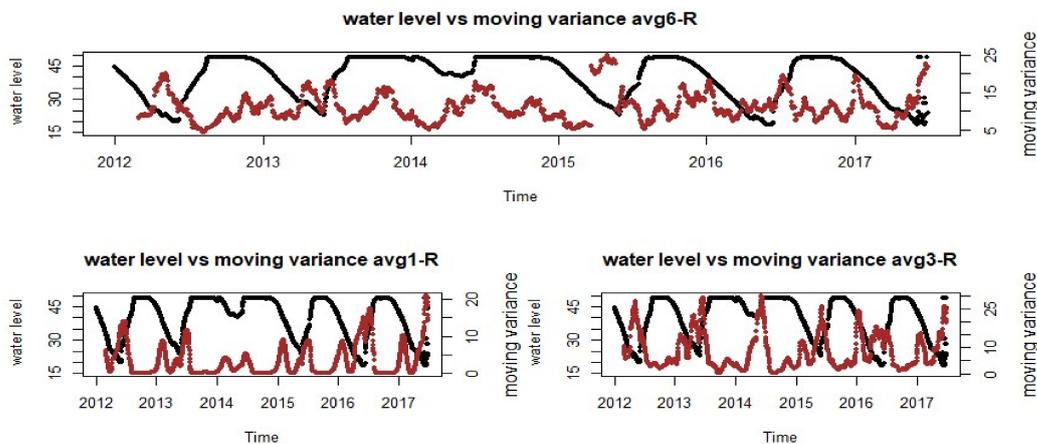

Figure 17. Avg6-R, avg1-R, and avg3-R 60-day moving variance (°C) (brown) against water level (m) (black).

Avg1-R and avg-2 R have a positive correlation with an almost y = x relationship (Figure 18); this applies to avg3 and avg4 (Figure 19). There is a slight difference between avg5-R and avg6-R, creating scatter (Figure 20). This behaviour is caused by the sun's trajectory, which is on the upstream side of the dam wall, with mountains on either side of the dam. Thermometers avg5 and avg6 are reached by the sun at different times of the day. This behaviour is not recorded by other thermometers, avg1–avg4, because of their locations on the dam wall.
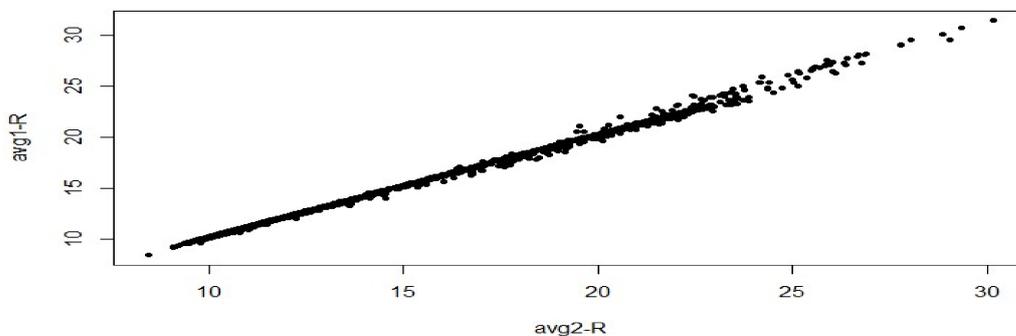

Figure 18. Avg1-R and avg2-R relationship.
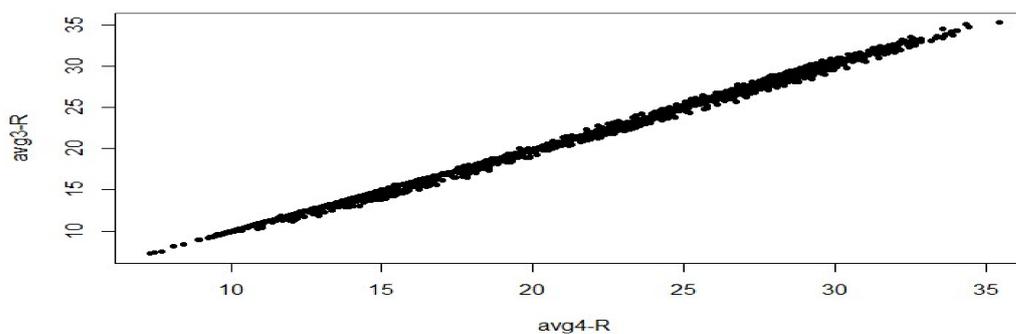
Figure 19. Avg3-R and avg4-R relationship.
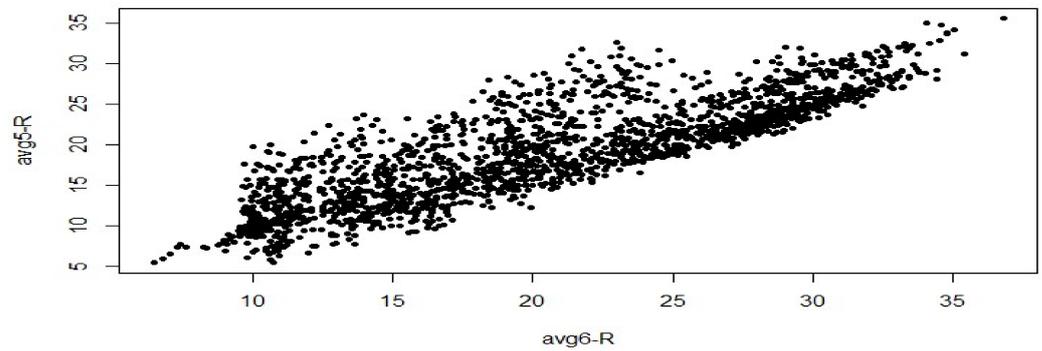
Figure 20. Avg5-R and avg6-R relationship.

The two main parameters that control the temperature of the water are the inflows and the ambient temperature. Figure 21 shows a relationship between ambient temperature and water level and between water level and avg6-R. The right plot on Figure 21 shows that water temperature decreases during dam filling and increases at full supply level until the following filling. Figure 22 shows the relationship between avg4-R and avg5-R, which follow the same pattern as that of avg6-R. Figure 23 shows the behaviours of avg2-R and avg3-R against water level.

Figure 21. Ambient temperature (°C) against water level (m) and water level (m) against avg6-R (°C).

Figure 22. Water level (m) against avg4-R and avg5-R water temperatures (°C).

Figure 23. Water level (m) against avg2-R and avg3-R water temperatures (°C).

Figure 23 shows the water level against avg2-R and avg3-R. Avg3-R scatters below 40 m, and avg2-R scatters below 25 m. This is caus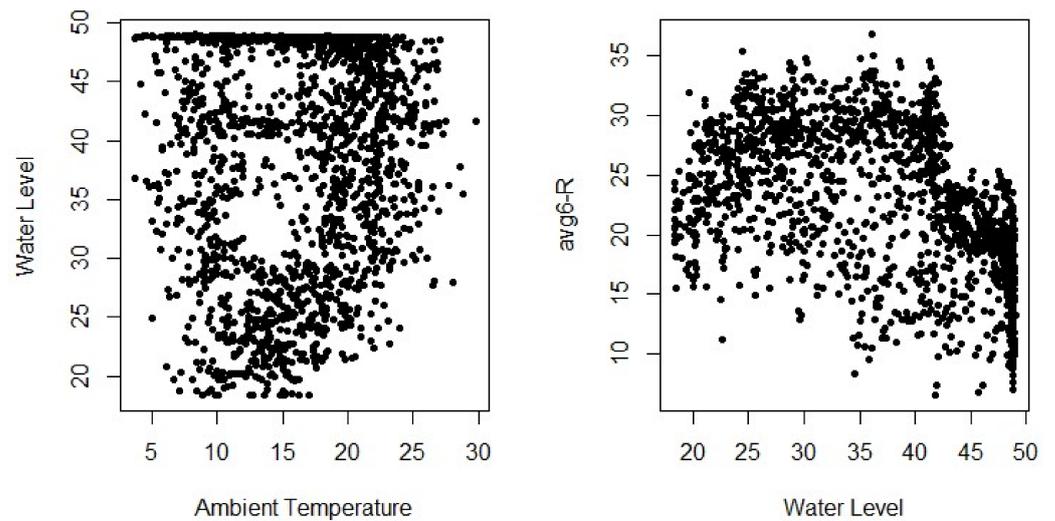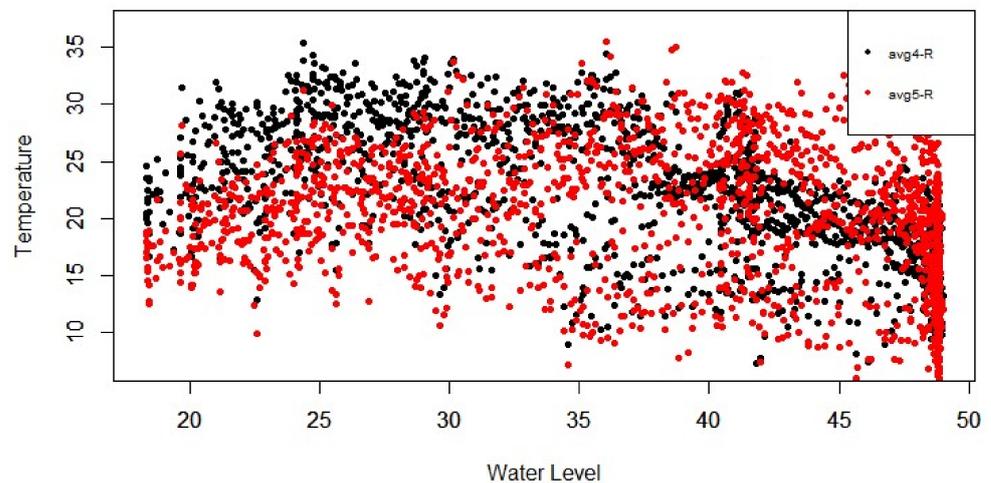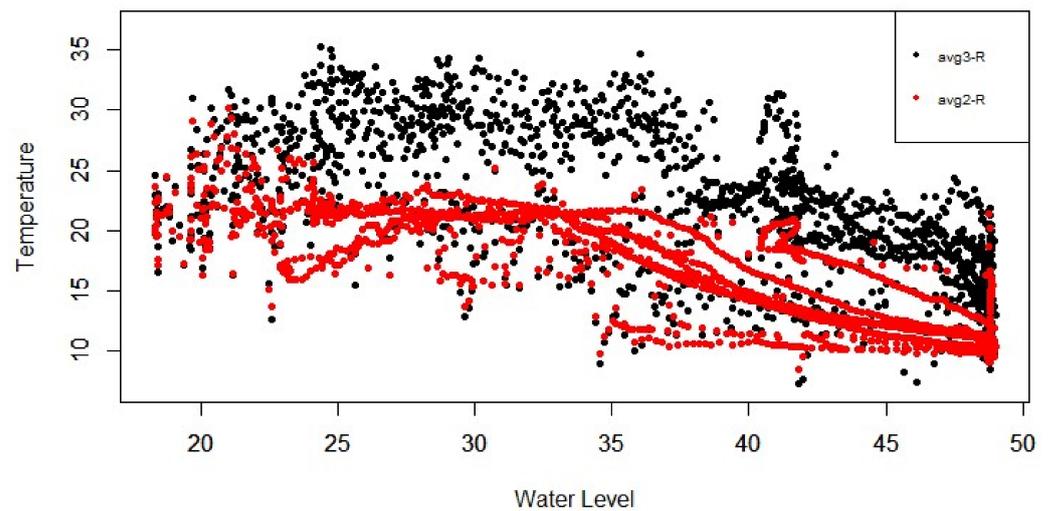ed by the locations of the thermometers avg2-R, which is at 26.23 m, and avg3-R, which is at 43.62 m. The data scatters when the water level is below the thermometer level; the scatter shows areas of high variance and where the variable has less influence on the behaviour of the dam with respect to the closure temperature.

Section 2.3 presents the machine-learning models that were used in this paper.

2.3. Nonparametric Machine-Learning Data-Model Selection

Machine-learning algorithms create models from training data for the purposes of estimation, prediction, and classification [14]. They are divided into parametric and nonparametric models. Parametric machine-learning data models assume a finite set of parameters θ. Given these parameters, future predictions, x, are independent of the observed data, D [15].

$$P(x|\theta, D) = P(x|\theta) \tag{1}$$

Everything there is to know about the data is captured by θ. Therefore, the complexity of the model is bounded even if the amount of data is unbounded, making parametric models not very flexible [16].

Nonparametric models assume that the data distribution cannot be defined in terms of a such a finite set of parameters but can be defined by assuming an infinite dimensional θ, usually thought as a function. The amount of information that θ captures about the data D can grow as the amount of data grows, making nonparametric data models more flexible [15].

That said, nonparametric machine-learning data models become suitable for evaluating the impacts of climate change permutations on the temperature loads of Roode Elsberg dam. There are two nonparametric machine learning data models that were studied for this document: Artificial Neural Networks (ANNs) and Random Forest (RF). Four algorithms were trained for ANN and RF and tuned accordingly; these included hybrid models. Hybrid models have been shown to possess an upper hand over other machine-learning data models [17]. Below is a brief description of how ANN and RF generally learn the data.

2.3.1. Artificial Neural Network

Artificial Neural networks (ANN) can mimic complex nonlinear relationships, and approximating by any measurable function [18], they have the capability to implement massive parallel computations for mapping, function approximation, classification, and pattern recognition [19].

An ANN is a biologically inspired computational model formed from hundreds of single units, artificial neurons, that relate to coefficients (weights) which constitute the neural structure [20]. Each neuron has weighed inputs, a transfer function, and one output, Figure 24. The neuron or processing element (PE) is essentially an equation which balances inputs and outputs [21]. A single neuron can compute simple information processing function, but the power of neural computations comes from connecting neurons in a network [20].
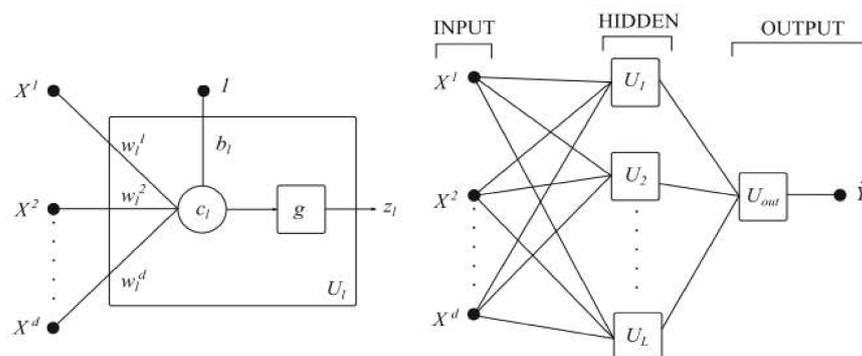
Figure 24. Left, schematic model of a perceptron U_1 and right, multilayer perceptron formed by L units, $U_1, \ldots, U_L$.

There are various types of ANN models [17], but vast applications for dam-monitoring data analysis are based on the multilayer perceptron [22]. In principle, several hidden layers can be used (Figure 9), but one is mostly adopted in practice [17].

The input of each unit U_l (Figure 24) is a linear combination of the predictors X^j:

$$c_l = \sum_{j=1}^{d} X \cdot w_1^j + b_l \tag{2}$$

which is transformed by an activation function g to compute the neuron's output:

$$z_l = g(c_l) \tag{3}$$

There are several forms of the activation function, g, that can be chosen, nonlinear in general. Figure 25 shows some of them. Sigmoid functions are often employed, such as the logistic [23], Mexico-hat, and the hyperbolic tangent [24].
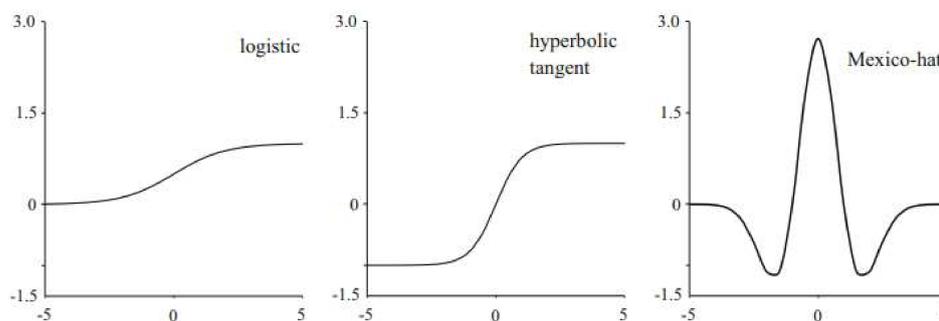
Figure 25. Common activation functions in Artificial Neural Networks (ANNs).

Depending on the activation function used in a neuron, the output layer (although a linear transform is frequently chosen), the overall model output is computed by the following [23]:

$$\hat{Y} = \sum_{l=1}^{L} w_{out}^l \cdot g\left(\sum_{j=1}^{d} X^j w_l^j + b_l\right) + b_{out} \tag{4}$$

ANNs can be thought of as multiple linear regression (MLR), for which output c_1 is expanded by the perceptron through a nonlinear transformation, g [25].

Sigmoid functions have a linear interval; thus, a unit with small weights performs a linear transformation, but they do exhibit horizontal asymptotes, which may cause numerical problems [26].

The most common learning algorithm is called back-propagation: ANN model parameters $\left\{ w_l^j, b_l, w_{out}^l, b_{out} \right\}$ are randomly initialized and iteratively undated to minimize a cost function (squared error sum, typically) by means of the gradient descent method [25].

The following algorithms were trained for the ANN:

- Model averaged neural network
- Neural network with feature extraction
- Bayesian regularized neural network
- Quantile regression neural network

2.3.2. Random Forest

Ordinary random forests grow an ensemble of trees, using n independent observations:

$$(Y_i X_i), i = 1, \ldots, n. \tag{5}$$

This algorithm grows a "forest", which has many trees, for which randomness is employed when selecting a variable at each tree split [27]. The size of the random subset, *mtry*, is the single tuning parameter, though depending on a specific random forest algorithm, there will be some optional parameters available for tuning. For regression, the prediction of random forest is the average of all trees.

The prediction of a single tree $T(\theta)$ for a new data point $X = x$ is obtained by averaging over the observed values on leaf $l(x, \theta)$. Let the weight vector $w_i(x, \theta)$ be given by a positive constant if the observation X_i is part of the leaf $l(x, \theta)$ and 0 if it is not. The sum of the weights is one, and therefore,

$$w_i(x, \theta) = \frac{1_{\{X_i \in R_{l(x,\theta)}\}}}{\#\left\{ j : X_j \in R_{l(x,\theta)} \right\}} \tag{6}$$

A single tree's prediction, which is given by covariate $X = x$, is the weighed average of the original observations $Y_i, i = 1, \ldots, n$:

$$\text{single tree}: \ \hat{u}(x) = \sum_{i=1}^n w_i(x, \theta) Y_i \tag{7}$$

using random forest, the conditional mean $E(Y|X = x)$ is approximated by the averaged prediction of k single trees, each constructed with an i.i.d. (independent and identically distributed) vector $\theta_t, t = 1, \ldots, k$. Let $w_i(x)$ be the average of $w_i(\theta)$ over this collection of trees:

$$w_i(x) = k^{-1} \sum_{t=1}^k w_i(x, \theta_t) \tag{8}$$

Then, an ordinary random forest prediction is given by

$$\text{Random Forests}: \ \hat{u}(x) = \sum_{t=1}^n w_i(x, \theta_t) \tag{9}$$

Ordinary random forest approximates the conditional mean $E(Y|X = x)$ by a weighed means over the observations of response Y. The conditional distribution function of Y, given that $X = x$, is given by

$$F(y|X = x) = P(Y \le y|X = x) = E\left(1_{\{Y \le y\}} \Big| X = x \right) \tag{10}$$

This expression is suited for drawing analogies with random forest approximation of the conditional mean. Just as $E(Y|X = x)$ is approximated by a weighed mean over

the observations of Y, approximation to $E\left(1_{\{Y \leq y\}} \middle| X = x\right)$ by the weighed mean over the observations of $1_{\{Y \leq y\}}$ is define by

$$F'(y|X = x) = \sum_{i=1}^{n} w_i\ (x) 1_{\{Y_i \leq y\}}, \tag{11}$$

The following algorithms were trained for RF:

- Ordinary RF
- Conditional inference RF
- Quantile RF
- Regularized RF

2.3.3. Machine-Learning Model Training

Before the water temperature and climate predictions data were used for training and testing of the machine learning algorithms, the following was done:

i. Data exploration
ii. Data cleaning
iii. Feature Engineering—the process of transforming raw data into features that better represent the underlying problem to the predictive models.
iv. Algorithm training with different hyperparameters for model selection—these hyperparameters were chosen at random using fit control.

2.3.4. Cross Validation

Cross validation was used during the model-training process, a method of obtaining a reliable estimate of model performance using only the training data. Only the training data were used during cross validation to reserve a truly exogenous dataset that was never used in training to test how well the model performs in the end. The are several ways in which data can be cross-validated, but repeated K-fold cross validation was chosen for this research. A 10-fold cross validation was used, with 5 repeats, Figure 26.
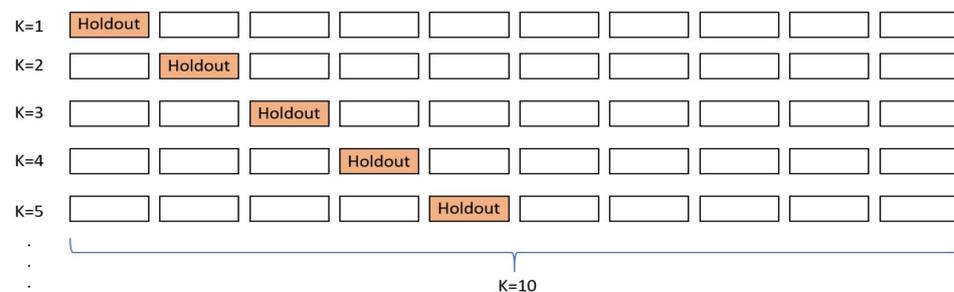
Figure 26. Model tuning through cross validation.

K-fold cross validation is robust in estimating the accuracy of a model, compared to others like leave-one-out cross validation. It gives more accurate estimates of the test error rate. Normally, performing k-folds using k = 5 or k = 10 has been shown empirically to yield test error rate estimates that suffer neither from excessively high bias nor from very high variance. K = 10 was used because of the size of the dataset.

ANN Cross Validation

Tables 1–4 show the ANN models that were cross-validated. The activation function across the four algorithms was kept constant, which is the logistic sigmoid function. The logistic sigmoid function was chosen because it is differentiable, meaning gradient-based backpropagation can be used with it, and most importantly because its outputs ranges between 0 and 1; therefore, it generates probabilities. During cross validation, the search for the best permutation of the hyperparameters was set at random.

Table 1. Model averaged neural network.

Size	Decay	Bag	RMSE	Rsquared	MAE
6	0.0080	FALSE	0.0196	0.0516	0.0154
16	0.7489	TRUE	0.0251	0.0130	0.0208
19	**0.0017**	**FALSE**	**0.0193**	**0.0531**	**0.0152**

Model averaged neural network used three hyperparameters for tuning: size (batch size), decay (weight decay), and bag (bagging).

Table 2. Neural network with feature extraction.

Size	Decay	RMSE	Rsquared	MAE
3	**0.0050**	**0.0195**	**0.8505**	**0.0152**
6	0.5848	0.0316	0.8340	0.0295
19	1.4074	0.0331	0.8310	0.0307

Neural network with feature extraction used two hyperparameters for tuning: size and decay.

Table 3. Bayesian regularized neural network.

Neurons	RMSE	Rsquared	MAE
7	0.0025	0.9700	0.0018
16	0.0022	0.9749	0.0016
17	**0.0022**	**0.9760**	**0.0016**

Bayesian regularized neural network used one hyperparameter for tuning: neuron.s.

Table 4. Quantile regression neural network.

n.hidden	RMSE	Rsquared	MAE
1	0.0037	0.9348	0.0028
3	0.0029	0.9586	0.0020
5	**0.0028**	**0.9587**	**0.0019**

Quantile regression neural network used one hyperparameter for tuning: n.hidden (number of hidden layers).

Table 5 shows a comparison between the four best-performing tunes out of the ANN machine-learning algorithms tuned.

Table 5. Summary of trained neural networks.

Algorithm	RMSE	Rsquared	MAE
Model averaged NN	0.0193	0.0531	0.0152
NN with feature Extraction	0.0195	0.8505	0.0152
Bayesian regularised NN	**0.0022**	**0.9760**	**0.0016**
Quantile Regression NN	0.0028	0.9587	0.0019

Random-Forest Cross Validation

Tables 6–9 shows the RF models that were tuned. For RF, just as for the case in ANN, during cross validation, the search for the best permutation of hyperparameters was set to random.

Table 6. Ordinary random forest.

min.node.size	mtry	splitrule	RMSE	Rsquared	MAE
6	**9**	**extratrees**	**0.0017**	**0.9860**	**0.0012**
16	7	maxstat	0.0019	0.9824	0.0013
19	6	extratrees	0.0020	0.9806	0.0015

Ordinary RF used three hyperparameters for tuning: min.node.size (minimum node size), mtry (number of variables randomly sampled), and split rule.

Table 7. Regularized random forest.

mtry	ConfReg	CoefImp	RMSE	Rsquared	MAE
1	0.745	0.756	0.0022	0.9761	0.0164
2	0.431	0.131	0.0019	0.9830	0.0014
7	**0.490**	**0.605**	**0.0016**	**0.9879**	**0.0011**

Regularized RF used three hyperparameters for tuning: mtry, ConfReg, and CoefImp.

Table 8. Conditional inference random forest.

mtry	RMSE	Rsquared	MAE
2	0.0026	0.9671	0.0020
5	0.0021	0.9788	0.0015
9	**0.0020**	**0.9810**	**0.0014**

Conditional inference RF used one hyperparameter for tuning: mtry.

Table 9. Quantile random forest.

mtry	RMSE	Rsquared	MAE
1	0.0020	0.9809	0.00133
7	**0.001547**	**0.9883**	**0.00107**
8	0.001548	0.9883	0.00108

Quantile RF used one hyperparameter for tuning: mtry.

Table 10 shows a comparison between the four best-performing tunes out of the RF machine-learning algorithms tuned.

Table 10. Summary of random forests trained.

Algorithm	RMSE	Rsquared	MAE
Ordinary Random Forest	0.0017	0.9860	0.0012
Conditional Inference RF	0.0020	0.9810	0.0014
Quantile RF	**0.0015**	**0.9883**	**0.0011**
Regularized RF	0.0016	0.9879	0.0011

From the models presented for ANN and RF, quantile random forests showed the best performance; therefore, it was used for this study.

2.3.5. Data Splitting

The water temperature dataset was imputed using random forest to fill any missing data; this was only done during the model selection phase. Then, the dataset was split into training and testing sets (Figure 27). Within this dataset, some data were spent on training

the models and some were spent on testing the models. The data that were reserved for testing were not used during training the model; this is because values of the dependant variable were predicted during the testing phase and the model accuracy was calculated based on the prediction error.

Figure 27. Data splitting.

3. Climate Change Model Data

Section 3 introduces the concept of climate change. A brief review of climate change models is given, Representative Concentration Pathways (RCPs) are explained, and the climate change models used in this study are shown.

Climate change models are used to investigate the response of the climate system to various forcings, to produce climate predictions on seasonal to decadal time scales, and to make projections of the future climate over the coming century and beyond [27]. These models are based on well-documented physical processes to simulate the transfer of energy and materials through a climate system by using mathematical equations to characterize how energy and matter interact in different parts of the ocean, atmosphere, and land [28].

Climate change models separate the earth's surface into three-dimensional grid cells. Each cell is modelled, and the results of the processes are passed to neighbouring cells to model the exchange of matter and energy over time [29]. These grid cells define the resolution of the model: the smaller the size of the grid cells, the more the level of detail in the model [30]. These models are validated using hind-casting tests; therefore, if the model performs well, its results for simulating future climate are also assumed to be valid [31]. To project future climate, the climate forcing is set to change according to a possible future scenario. Scenarios are possible permutations of how quickly human population will grow, land will be used, economies will evolve, and atmospheric conditions will be a result of each permutation [32].

These scenarios, also called Representative Concentration Pathways (RCPs), are used by the Intergovernmental Panel on climate Change (IPCC), [33]. These pathways are characterized by the radiative forcing produced by the end of the 21st century [34]. Radiative forcing is the extra heat the lower atmosphere will retain as a result of additional greenhouse gases [35].

The following are RCPs as recognized by the IPCC:

i.   RCP 2.6—this pathway sees emissions that peak early and then fall due to active removal of atmospheric carbon dioxide.
ii.   RCP 4.5—this pathway stabilizes total radiative forcing to 4.5 W/m^2 shortly after 2100 by the application of a range of technologies and strategies for reducing greenhouse emissions.
iii.   RCP 6.0—this pathway stabilizes total radiative forcing to 6 W/m^2 shortly after 2100 by the application of a range of technologies and strategies for reducing greenhouse emissions.
iv.   RCP 8.5—this pathway has its forcing pathways rising to 8.5 W/m^2 by 2100.

The South African climate change atlas is based on RCP 4.5 and RCP 8.5 [35].

Climate model complexity has grown over time because of the additional incorporation of components of earth's climate system. Climate change today can simulate many aspects of the climate system like the atmospheric chemistry and aerosols; land surface in-

teractions including soil and vegetation, land, and ice; and increasingly even an interactive carbon cycle and/or biogeochemistry [36]. The number of climate change models have also increased in the modern day, as computers become powerful, and with each successive version of the World Climate Research Programme's (WCRP's) Coupled Model Interpolation Project (CIMP) [37]. CIMP5 provides output from over 50 GCMs (General Circulation Models), with spatial resolutions ranging from about 50 km–300 km per horizontal size and variable vertical resolution on the order of hundreds of meters in the troposphere or lower atmosphere [38], and prediction permutations from these 50 models are equally possible [38].

The assumption is that higher-resolution, more complex, and more up-to-date models (CIMP5) will produce more robust projections than previous generation models (CIMP3). However, research comparing CIMP3 and CIMP5 simulations concluded that, although the spatial resolution of CIMP5 has improved relative to CIMP3, the overall improvement in performance is relatively minor [37]. Though CIMP5 simulations do show modest improvement in the model ability to simulate some aspects of cloud characteristics [37] and the rate of arctic sea ice loss [39], CIMP5 models are proposed for this research.

3.1. Climate Change in South Africa

Figure 28 shows the climatic conditions is South Africa. South Africa's climate conditions generally range from temperate coastal in the southwest corner of the country to temperate interior in the plateau and hot in the interior in the northeast. The country's climate is generally warm, with sunny days and cool nights, and rainfall occurs in summer and autumn (November–March), with Cape Town experiencing winter rainfall (June–August). Temperatures are mostly influenced by variations in elevation, terrain, and ocean currents as opposed to latitudes [40]. The climatic conditions vary in response to movement of the high-pressure belt that circles the globe between 25°S latitude and 30°S latitude during winter and low-pressure systems that occur during winter [40]. There are very little temperature differences in terms of average from the south to north; the variation can be noticed though between the east and west [34].

South Africa, like many other developing countries, is especially vulnerable to the impacts of climate change. Precipitation is the primary medium through which the impacts of climate change are felt in South Africa, according to the National Water Resource Strategy [33].

The following climate change scenarios are expected in South Africa:

- Warmer (<3 °C above 1961–2000) and wetter, with a greater frequency of extreme events
- Warmer (<3 °C above 1961–2000) and drier, with an increased frequency of drought events and somewhat greater frequency of extreme rainfall events
- Hotter (>3 °C above 1961–2000) and drier, with a substantial increase in the frequency of drought events and greater frequency of extreme rainfall events
- Hotter (<3 °C above 1961–2000) and drier, with a substantial increase in the frequency of drought events and greater frequency of extreme events
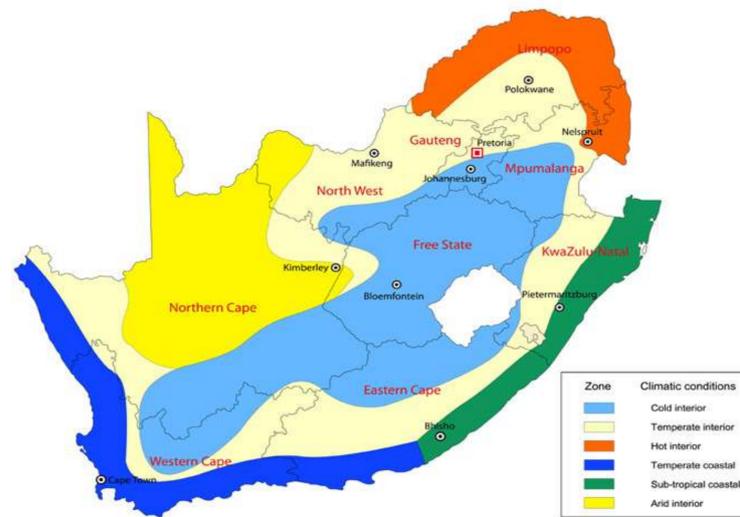
Figure 28. Climatic conditions in South Africa [41].

3.2. Yearly Predicted Ambient Temperatures at RCP 4.5 and RCP 8.5

The South African climate atlas concentrates on RCP 4.5 and RCP 8.5; therefore, these are the pathways that were used in this paper. Data shown in Figures 29 and 30 were downloaded from https://esgf-node.llnl.gov/projects/esgf-llnl/; they are from the CIMP5 as daily readings. A total of 18 models were downloaded for both RCP 4.5 and RCP 8.5; these models were plotted at each RCP, and the hottest model (hot model) and the least hot model (warm model) were chosen.
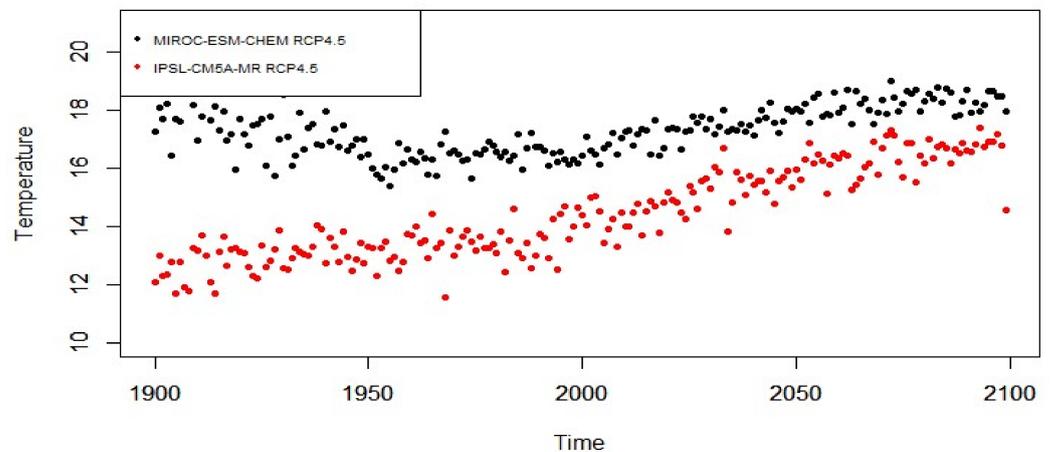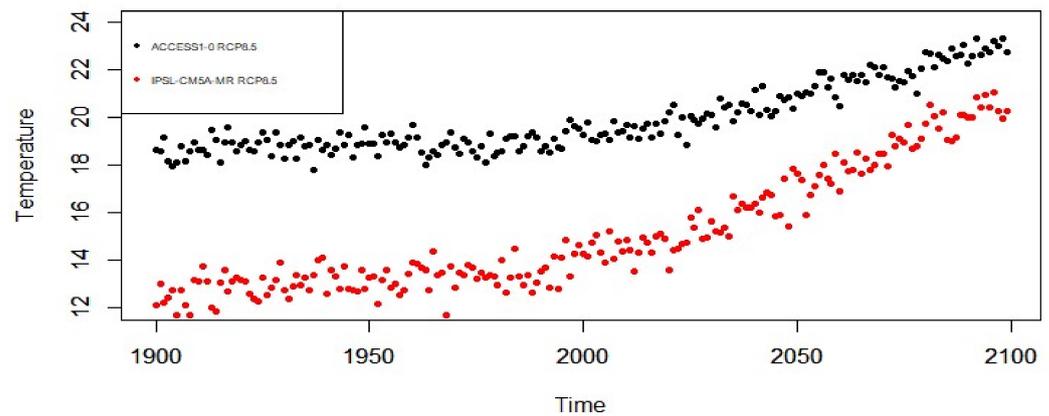

Figure 29. Representative Concentration Pathway (RCP) 4.5 hot model (MIROC-ESM-CHEM) and warm model (IPSL-CM5A-MR), temperature in °C.

Figure 30. RCP 8.5 hot model (ACCESS1-0) and warm model (IPSL-CM5A-MR), temperature in °C.

The hot model and warm model were used at each RCP to account for uncertainties inherent in the climate change model's data, thus giving lower and upper temperature bounds. This gave a total of eight climate permutations studied in this document. These four models presented four temperature variables that were studied. The choice of models did not include the amplitude of the climate models.

3.3. Roode Elsberg Dam Water-Level Scenarios

Two water-level scenarios were studied for this paper:

- A case where the recorded water level was used—for this case, the water level was kept constant for periods 2012–2017 and 2048–2053, reasons for the choice explained in Section 3.3.1.
- The second case is for a full dam steady-state, which was also kept constant for periods 2012–2017 and 2048–2053, explained in Section 3.3.2.

3.3.1. Recorded Water Level 2012–2017

Rainfall is another variable used to determine the permutations used. Figure 31 below shows the annual rainfall of Klondyke farm, which is within Roode Elsberg's catchment, against the model that predicts the wettest rainfall for the region at RCP 8.5, which is the worst case. The figure shows a reducing trend of rainfall in the region. Water levels are affected by inflows and outflows; inflows control how quickly the dam fills up, and ouflows control how long the dam stays full. Roode Elsberg is used for irrigation; therefore, the outflows will be controlled by water demand, which has been shown to have a possibility of increasing due to climate change [34]. Inflows are affected by precipitation; therefore, with rainfall of the region decreasing, it can be assumed that the rate of filling of the dam will not change and that, with projected increase in water demand, the period when the dam will be full is likely to reduce. Given the above assumptions, the recorded water level (Figure 13) became a constant during training and making predictions.
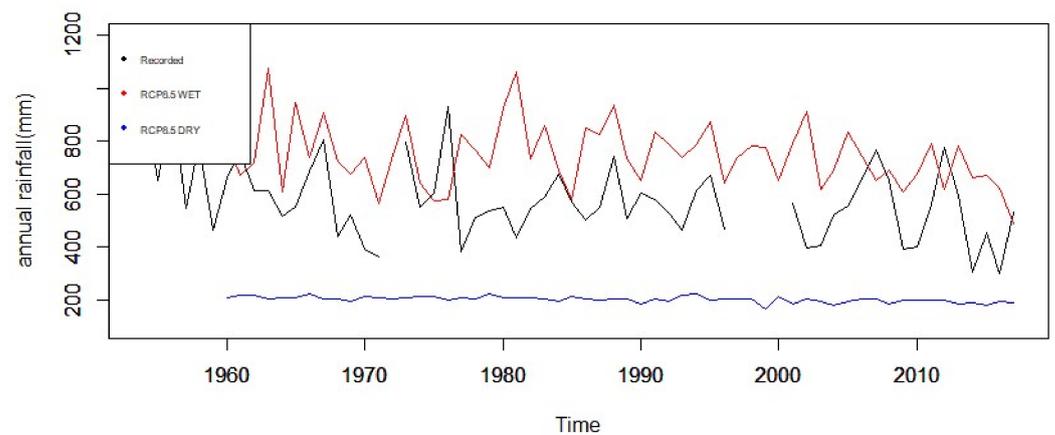
Figure 31. Klondyke farm annual rainfall.

3.3.2. Roode Elsberg Dam Full Dam Steady-State

Water level is also limited by the dam height; therefore, a special case (steady-state case) of when the dam was full was created using a 2014–2015 wet year. When Roode Elsberg loses most of its water, the dam fills up on average beginning in august and the water decreases around the end of November, with irrigation of vineyards at the peak. During the wet year, the dam filled up around mid-June. Therefore, when creating the steady-state, the above conditions were accounted for. Figure 32 shows the steady-state.
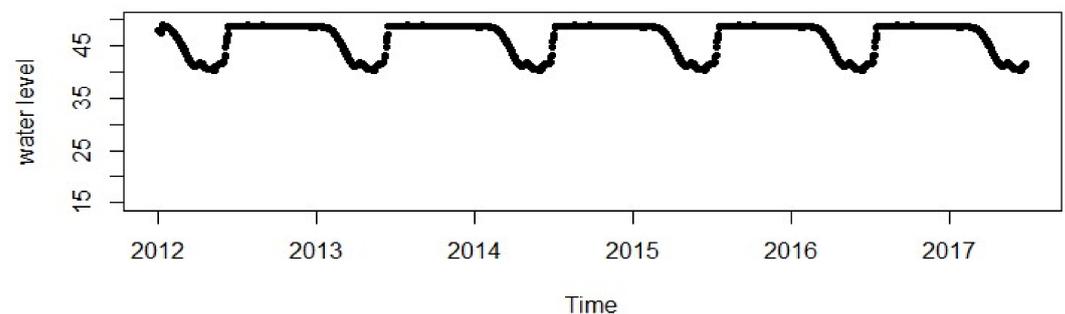

Figure 32. Full dam steady-state, water level in m.

For the relationships between the variables for water temperature, water level, and ambient temperature, shown by Figures 8–23, as well as the assumptions made in relation to Roode Elsberg's water level, the following methodology was designed. The training period was from January 2012–June 2017; this was controlled by the water level data for the same period.

4. Methodology for Predicting Water Temperatures

The following methodology was designed with an aim to predict water temperatures for Roode Elsberg dam at RCP 4.5 (hot and warm model) and RCP 8.5 (hot and warm model). This created a total of 8 climate change permutations at RCP 4.5 and 8 climate permutations at RCP 8.5 for periods 2012–2017 and 2048–2053, accounting for water level and ambient temperature variable changes. Predictions were made for the period 2012–2017 and 2048–2053 (mid-century).

Figure 33 shows the 6 models that were trained for predicting different water temperatures and the variables used to train each model. This also involved variables used to make predictions. There are six models that were trained for each water temperature, and respective variables were used. The choice of variables was determined using the relationships between the variables and the transfer of thermal energy within the dam.
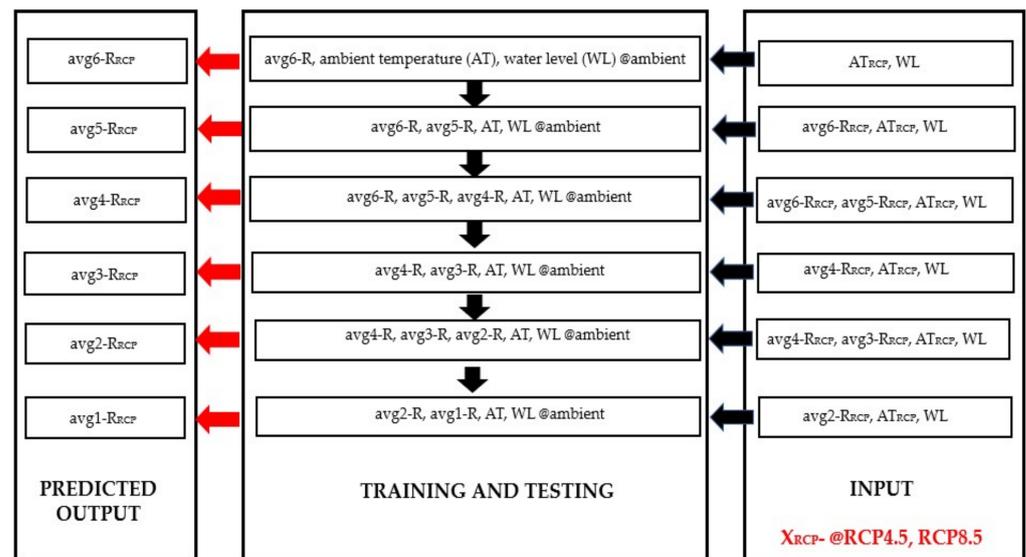
Figure 33. Methodology for predicting water temperatures.

5. Results

For this study, there were two variables that were studied: ambient temperature and water level. Ambient temperature was changed using climate change models at RCP 4.5 and RCP 8.5. There were two models used at each RCP: hot and warm models. These are shown by Figures 29 and 30, and there were two water-level cases used as explained in Sections 3.3.1 and 3.3.2.

The recorded water temperatures from 2012–2017 was used for training and testing the quantile RF model, and then, ambient temperature and water level variables were changed to make predictions for the periods 2012–2017 and 2048–2053. The thermometers on the right flank of the dam were used for the analysis.

5.1. Quantile Random Forest Training

Six models were trained to predict different water temperature variables, as shown in Table 11 one for each water temperature variable. The models trained to predict the water temperatures from avg6-R to avg1-R showed an increase in terms of accuracy. The accuracy of each model was dependant on the relationship between the variables used to train the model.

Table 11. Predicting accuracy for water temperature predictions.

Thermometer	RMSE	Rsquared	MAE
avg6-R	1.868	0.892	1.372
avg5-R	1.930	0.927	1.381
avg4-R	1.434	0.960	0.941
avg3-R	0.424	0.996	0.221
avg2-R	1.154	0.944	0.650
avg1-R	0.230	0.998	0.093

5.2. Period 2012–2017 Predictions

The only variable that changed was ambient temperature. Predictions were made at both RCPs. The differences were too small to be analysed monthly and still be visible; therefore, yearly averages of the data were made in order to analyse the differences between the predictions.

5.2.1. Recorded Water Level

Figures 34 and 35 show temperature predictions for a warm model and a hot model at RCP 4.5 and RCP 8.5 respectively from 2012–2017 (i.e., avg6 (recorded data) and Pavg6

(predicted data). The RCP 4.5 yearly averages for the hot (MIROC-ESM-CHEM) and warm (IPSL-CHEM-MR) models were almost the same, and for RCP 8.5, the hot model (ACCESS1-0) predicted higher values for avg6-R, avg3-R, and avg1-R for the period 2012–2017 than the warm model (IPSL-CHEM-MR). Predictions at bo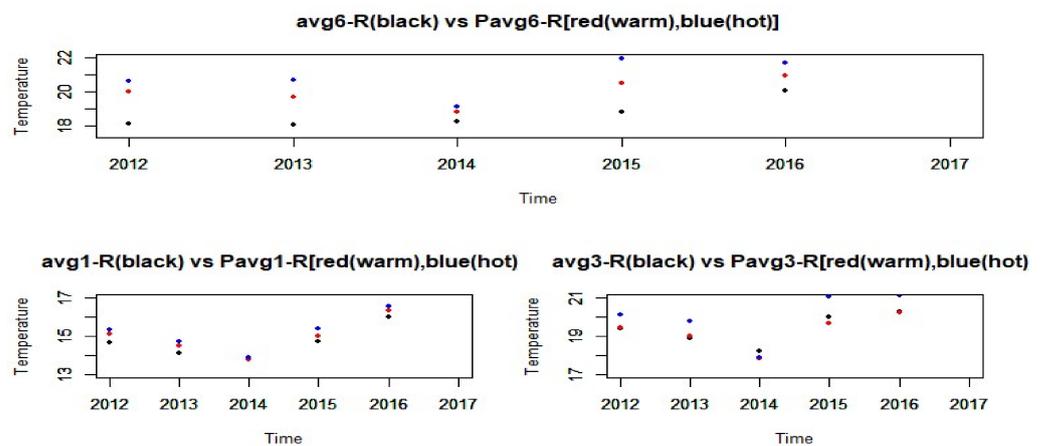th RCPs showed higher temperatures than the recorded water temperatures for the same period, therefore posing a question: Where does Roode Elsberg's climate fall on the RCP scale?
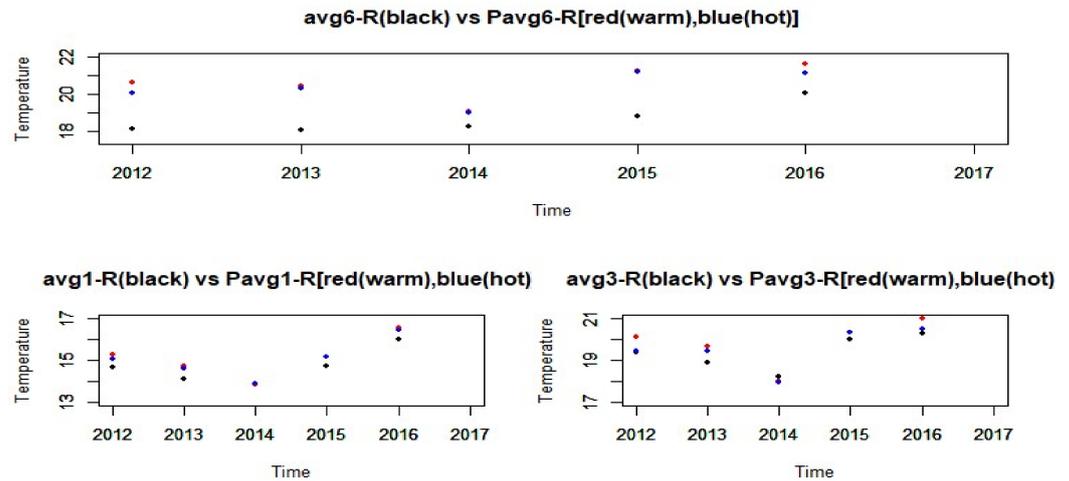
Figure 34. RCP 4.5 warm model (IPSL-CHEM-MR) and RCP 4.5 hot model (MIROC-ESM-CHEM) predictions for the period 2012–2017 (recorded water level), temperature in °C.

Figure 35. RCP 8.5 hot model (MIROC-ESM-CHEM) and RCP 8.5 warm model (IPSL-CM5A-MR) predictions for the period 2012–2017 (recorded water level), temperature in °C.

5.2.2. Full Dam Steady-State

Predictions were made at both RCPs, and yearly averages of the data were made in order to analyse the differences between the predictions. Figures 36 and 37 show temperature predictions for a warm model and a hot model at RCP 4.5 and RCP 8.5, respectively, from 2012–2017. The RCP 4.5 yearly averages for hot (MIROC-ESM-CHEM) and warm (IPSL-CHEM-MR) models had clear differences; this applied to RCP 8.5.

Figure 36. RCP 4.5 warm model (IPSL-CHEM-MR) and RCP 4.5 hot model (MIROC-ESM-CHEM) predictions for the period 2012–2017 (full dam steady-state), temperature in °C.

5.3. Period 2048–2053 Predictions

5.3.1. Recorded Water Level

Figures 38 and 39 show temperature predictions for a warm model and a hot model at RCP 4.5 and RCP 8.5, respectively, from 2048–2053. The RCP 4.5 yearly averages for a hot model (MIROC-ESM-CHEM) were higher than that of the warm model (IPSL-CHEM-MR) for the period, and for RCP 8.5, the hot model (ACCESS1-0) predicted higher values for avg6-R, avg3-R, and avg1-R for the warm model (IPSL-CHEM-MR).

Figure 37. RCP 8.5 hot model (MIROC-ESM-CHEM) and RCP 8.5 warm model (IPSL-CM5A-MR) predictions for the period 2012–2017 (full dam steady-state), temperature in °C.
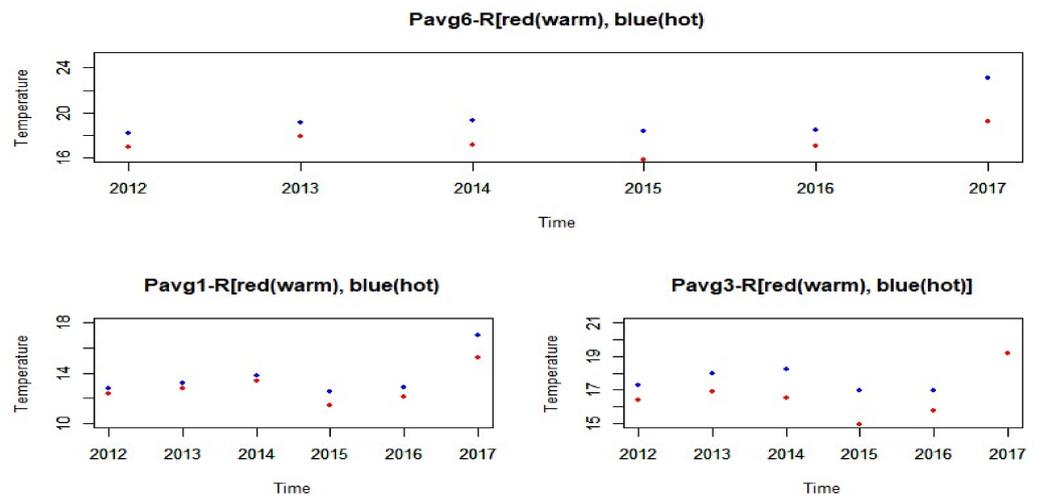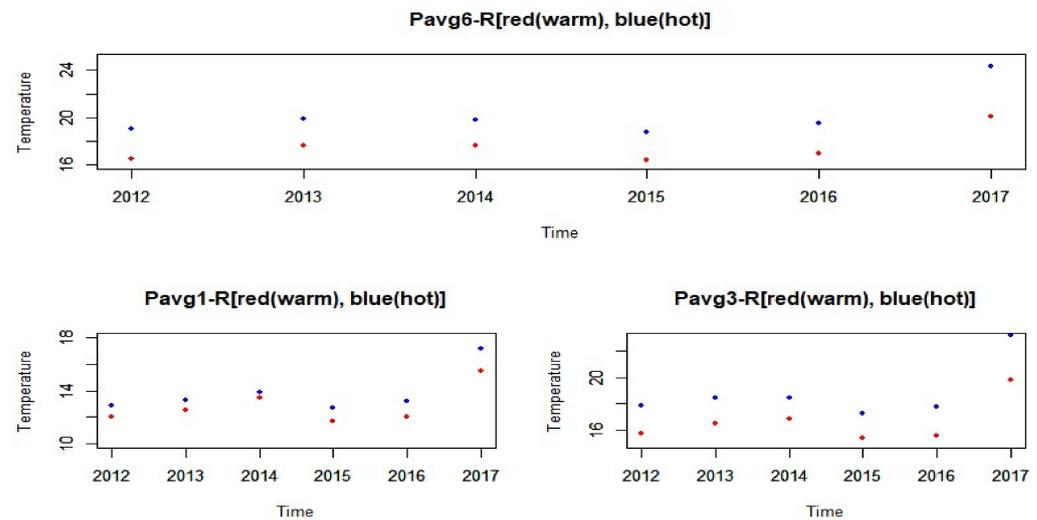
Figure 38. RCP 4.5 warm model (IPSL-CHEM-MR) and RCP 4.5 hot model (MIROC-ESM-CHEM) predictions for the period 2048–2053 (recorded water level), temperature in °C.
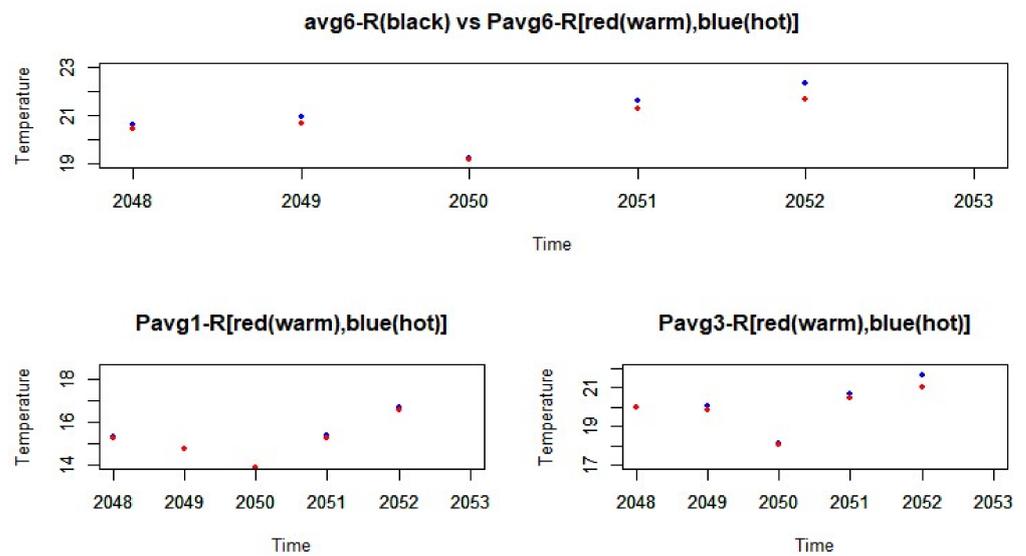

Figure 39. RCP 8.5 warm model (IPSL-CM5A-MR) and RCP 8.5 hot model (ACCESS1-0) predictions for the period 2048–2053 (recorded water level), temperature in °C.
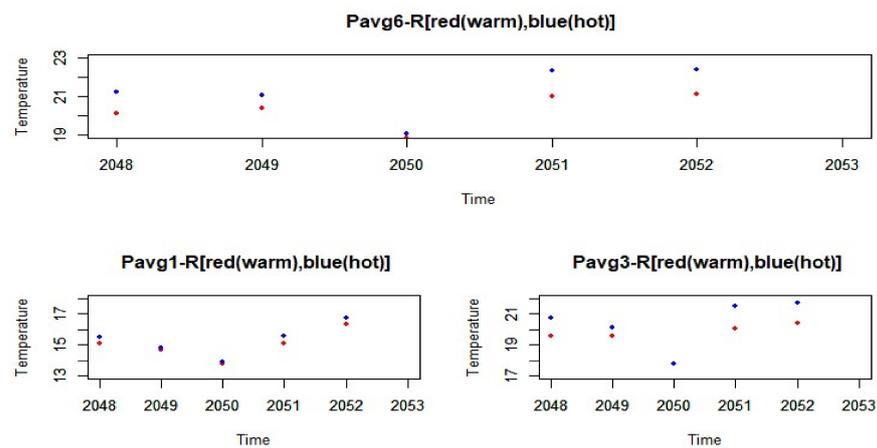
5.3.2. Full Dam Steady-State

Figures 40 and 41 shows temperature predictions for a warm model and a hot model at RCP 4.5 and RCP 8.5, respectively, from 2048–2053 for a full dam steady-state. The RCP 4.5 yearly averages for a hot model (MIROC-ESM-CHEM) were higher than that of the warm model (IPSL-CHEM-MR) for the period, and for RCP 8.5, the hot model (ACCESS1-0) predicted higher values for avg6-R, avg3-R, and avg1-R for the warm model (IPSL-CHEM-MR).

Figure 40. RCP 4.5 hot model (MIROC-ESM-CHEM) and RCP 4.5 warm model (IPSL-CM5A-MR) predictions for the period 2048–2053 (full dam steady-state), temperature in °C.
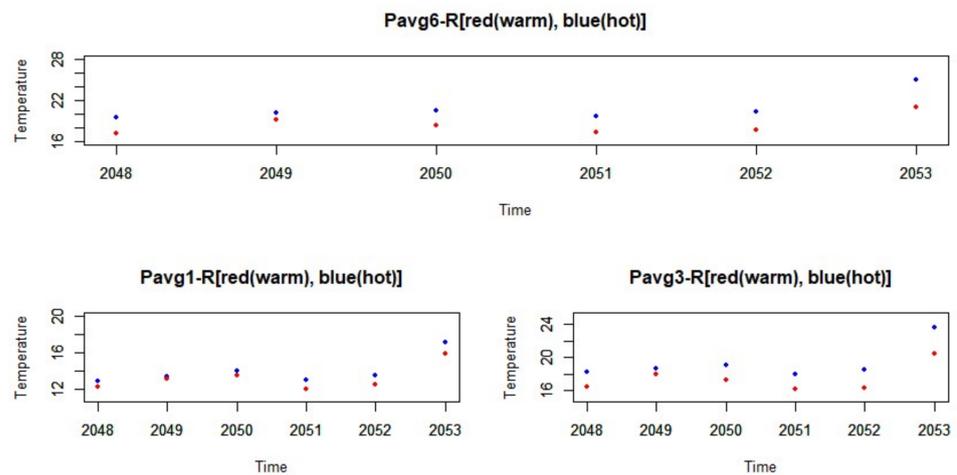
Figure 41. RCP 8.5 hot model (MIROC-ESM-CHEM) and RCP 8.5 warm model (IPSL-CM5A-MR) predictions for the period 2048–2053 (full dam steady-state), temperature in °C.
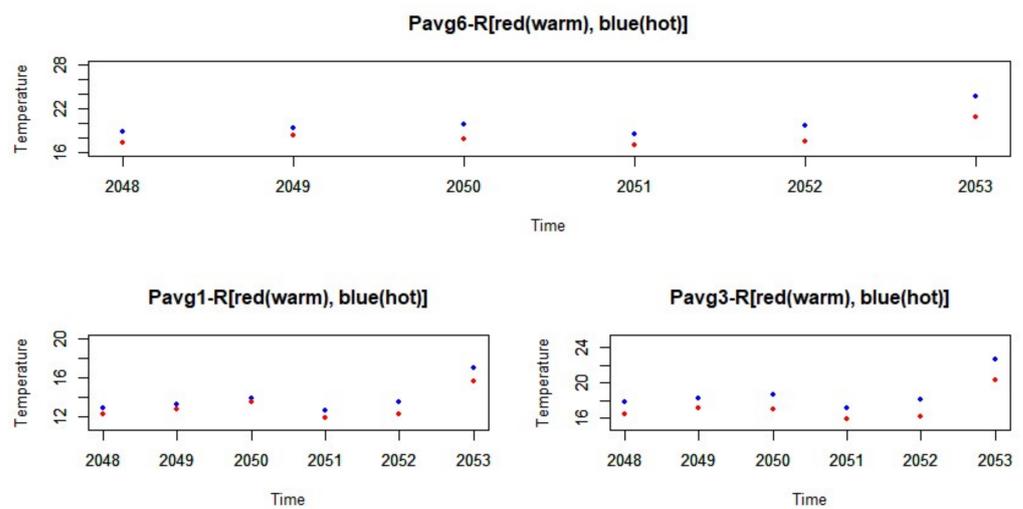
5.4. Average Temperatures 2012–2017

5.4.1. Recorded Water Level

Table 12 shows the average temperature for 2012–2017 for a case where recorded water level was used.

Table 12. Predicted average temperatures for 2012–2017, recorded water level.

Dataset	AT	avg6-R	avg3-R	avg1-R
recorded data	**15.9**	**20.9**	**19.3**	**15.0**
RCP 4.5 IPSL-CHEM-MR (Warm)	14.5	19.6	18.9	14.8
RCP 4.5 MIROC-ESM-CHEM (Hot)	17.0	21.6	20.7	15.7
RCP 8.5 IPSL-CHEM-MR (Warm)	14.4	19.6	19.0	14.8
RCP 8.5 ACCESS1-0 (Hot)	19.4	22.6	21.3	15.8

5.4.2. Full Dam Steady-State

Table 13 shows the average temperature for the period 2012–2017 for a case where the full dam steady-state was used for water level.

Table 13. Predicted average temperatures for 2012–2017, full dam steady state.

Dataset	AT	avg6-R	avg3-R	avg1-R
RCP 4.5 IPSL-CHEM-MR (Warm)	14.5	17.0	16.1	12.4
RCP 4.5 MIROC-ESM-CHEM (Hot)	17.0	18.6	17.5	13.0
RCP 8.5 IPSL-CHEM-MR (Warm)	14.4	17.0	16.0	12.4
RCP 8.5 ACCESS1-0 (Hot)	19.4	19.4	18.0	13.2

5.5. Average Temperatures 2048–2053

Tables 14 and 15 show the average temperatures for the period 2048–2053 for the cases where recorded water level was used and where full dam steady-state was used, respectively.

Table 14. Predicted average temperatures for 2048–2053, recorded water level.

Dataset	AT	avg6-R	avg3-R	avg1-R
RCP 4.5 IPSL-CHEM-MR (Warm)	15.9	20.4	19.5	15.0
RCP 4.5 MIROC-ESM-CHEM (Hot)	18.1	22.3	21.4	16.0
RCP 8.5 IPSL-CHEM-MR (Warm)	16.9	20.9	19.9	15.1
RCP 8.5 ACCESS1-0 (Hot)	20.9	23.3	22.0	16.1

Table 15. Predicted average temperatures for 2048–2053, full dam steady state.

Dataset	AT	avg6-R	avg3-R	avg1-R
RCP 4.5 IPSL-CHEM-MR (Warm)	15.9	17.6	16.5	12.6
RCP 4.5 MIROC-ESM-CHEM (Hot)	18.1	19.3	18.0	13.3
RCP 8.5 IPSL-CHEM-MR (Warm)	16.9	18.0	16.8	12.7
RCP 8.5 ACCESS1-0 (Hot)	20.9	20.0	18.5	13.4

5.5.1. Recorded Water Level

Table 14 shows the average temperatures for period 2048-2053 for a case where recorded water level was used.

5.5.2. Full Dam Steady-State

Table 15 shows the average temperature for period 20148-2053 for a case where full dam steady state was used.

5.6. Average Temperature Changes 2012–2053

Tables 16 and 17 show the effects of ambient temperature on water temperatures for Roode Elsberg dam from 2012–2053 at recorded temperature and full dam steady-state, respectively. An increase in ambient temperature leads to an increase in water temperature, though average water temperatures at recorded water levels remains higher than when the full dam steady-state was used. Average temperature increases for both water level cases differ slightly, on average by 0.2 °C. There is an average increase of 1.6 °C in the ambient temperature between 2012 and 2053. This increase has led to increases of 0.9 °C for avg6-R, 0.8 °C for avg3-R, and 0.4 °C for avg1-R at the recorded water temperature; the same increase in temperature led to increases of 0.7 °C for avg6-R, 0.6 °C for avg3-R, and 0.3 °C for avg1-R for a full dam steady-state.

Table 16. Effects of ambient temperature on water temperature for Roode Elsberg dam from 2012–2053 at different RCPs, recorded water level.

Dataset	AT	avg6-R	avg3-R	avg1-R
RCP 4.5 IPSL-CHEM-MR (warm)	1.4	0.8	0.6	0.2
RCP 4.5 MIROC-ESM-CHEM (Hot)	1.0	0.7	0.7	0.5
RCP 8.5 IPSL-CHEM-MR (warm)	2.6	1.3	1.0	0.4
RCP 8.5 ACCESS1-0 (Hot)	1.4	0.7	0.7	0.3

Table 17. Effects of ambient temperature on water temperature for Roode Elsberg dam from 2012–2053 at different RCPs, full dam steady state.

Dataset	AT	avg6-R	avg3-R	avg1-R
RCP 4.5 IPSL-CHEM-MR (warm)	1.4	0.6	0.4	0.2
RCP 4.5 MIROC-ESM-CHEM (Hot)	1.0	0.7	0.5	0.3
RCP 8.5 IPSL-CHEM-MR (warm)	2.6	1.0	0.8	0.3
RCP 8.5 ACCESS1-0 (Hot)	1.4	0.6	0.5	0.2

5.6.1. Recorded Water Level

Table 16 shows temperature change for a case where recorded water level was used for period 2012–2053.

5.6.2. Full Dam Steady-State

Table 17 shows temperature change for a case where full dam steady state was used for period 2012–2053.

5.7. Moving Water Temperature Average

Figure 42 below shows the moving average for predicted water temerarture using the recorded water level (black) and full dam steady-state (grey). The figure shows a decrease in the maximum temperatures of the moving average for the full dam steady-state. The full dam steady-state peaks later than the recorded water temperature case. It takes on average 53 days for the dam to feel the effects of temperature increase in the season when the dam lost most of its water, and it takes 79 days on average for the day to feel the effects of increase in temperature during a full dam's steady-state. This behaviour was also notice on the recorded data in the year 2014, when the dam did not lose a lot of water.
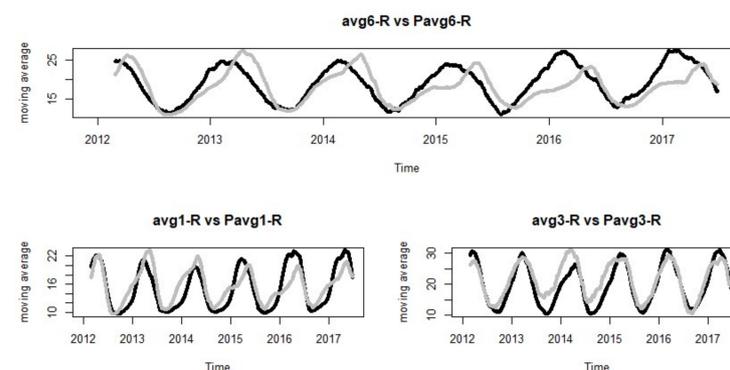
Figure 42. Moving averages (°C) for predicted water temperatures at the recorded water level (black) and at full dam steady-state (grey).

6. Discussion

Yearly average temperatures when recorded water level was used remained higher than yearly average temperatures when the full dam steady-state was used. This is very interesting because there are seasons when the water level was low on the recorded water

temperature, which is not studied for this paper. Does this mean that the dam wall gets exposed to more temperatures during seasons when the water level is low? It is definitely an area to explore, but the evidence from the data shows that it takes on average 53 days for the dam to feel the effects of temperature increase in the season when the dam lost most of its water and it takes 79 days on average for the day to feel the effects of an increase in temperature during a full dam's steady-state. Though the yearly average water temperatures when the recorded water level was used are higher than when the full dam steady-state was used, the temperature increases for the two cases did not differ much.

The increase in ambient temperature from 2012–2053 led to an increase in water temperatures, which increased with the height of the dam wall. An average temperature increase of 1.6 °C for ambient temperature led to average water temperature increases of 0.9 °C for avg6-R, which is at level 47.3 m; of 0.8 °C for avg3-R at level 43.62 m; and of 0.4 °C for avg1-R, which is at level 26.23 m. The same increase in average temperature led to average increases of 0.7 °C for avg6-R, 0.6 °C for avg3-R, and 0.3 °C for avg1-R for a full dam steady-state. This shows that, on average, the water temperature changes due to climate change do not vary much (they vary by 0.2 °C, on average) relative to climate models.

The two water-level cases show that average temperature changes between 2012–2053 are almost the same. Though this is the case, the mechanical behaviour of the dam in both cases is completely different. The dam exposed to the recorded water-level case experienced higher average temperature gradients than the dam in the full dam steady-state. The other aspect that is of importance is the difference in their 60-day water temperature moving averages (Figure 40). For the case where recorded water level was used, the water temperature moving averages increased with time while the water temperature moving averages decreased over time with a full dam steady-state.

The effects of water temperature are more important than the effect of ambient temperature in temperature distribution of the dam across the thickness, emphasising how critical water temperatures are to the behaviour of arch dams [14].

Figure 15 shows the 60-day moving averages for avg6-R, avg1-R, and avg3-R with respect to the closure temperature of the dam (horizontal blue line). The important aspect of this figure is the peak temperatures of the moving averages. In 2014, Roode Elsberg experienced a "full dam state", which was used to create the full dam steady-state. During this season, it be seen that the water temperatures peaked at different times. When the water level of the dam was low, as in any other year, avg3-R and avg1-R peaked around the same time while avg6-R peaked early. This highlights how, at respective times and for different cases, water level depths have different influences on the behaviour of the dam.

This is further illustrated by Figure 16, where the 60-day moving average was plotted against cumulative heat with respect to 1 January 2012. On Figure 16, the blue horizonal lines shows the closure temperature of the dam, which is approximately 15.5 °C. The figure shows that, for most of the year, the dam is heated (expanding) more than cooled (contracting); this varies at different depths of the dam. Avg6-R, avg1-R, and avg3-R have amplitudes on average of 7.5 °C, 10 °C, and 15 °C, respectively, above the closure temperature. Therefore, a rise in temperature will increase the amplitude of the 60-day moving average, meaning that the dam gets heated more than cooled. This shows that different levels of the dam experience different stresses with respect to time; a more robust analysis of this behaviour can be done on finite element modelling to assess this further.

According to [42], Roode Elsberg's left flank rock foundation has an elastic modulus of 25 GPa and the rock foundation on the right flank has an elastic modulus of 20 GPa. Therefore, the left and the right flanks of Roode Elsberg dam will expand and contract at different rates with respect to the foundation because of its properties (in the case of the Roode Elsberg foundation, monitoring data are not available). The projected increase in water temperatures will increase the thermal stresses of the dam and, because of the foundation properties, will influence the orientation of the dam. An increase in ambient temperature due to climate change will cause an increase in the water temperature and

hence thermal stresses. This will cause the dam to push upstream on the left flank more than on the right flank.

If both flanks push upstream at the same rate as a result of thermal stresses, the dam will experience in shift in its deformations. In the case of the Roode Elsberg dam, the rate at which the left flank pushes upstream will be more that the rate at which the right flank pushes upstream. The dam will experience two movements: a shift in its deformations and a change in the reference of GPS measurements caused by rotation. For Roode Elsberg, its deformations are referenced due north; as a result, the orientation will shift towards north-northwest as a result of increase in water temperatures.

7. Conclusions

Based on the study carried out, the following can be concluded:

- An increase in ambient temperature leads to an increase in the water temperatures of Roode Elsberg dam.
- The effects of climate change on water temperature are different at different depths of the dam; an average increase of 1.6 °C in the ambient temperature between 2012 and 2053 led to average increases of 0.9 °C, 0.8 °C, and 0.4 °C for avg6-R, avg3-R, and avg1-R, respectively, for recorded water levels, and the same increase in average temperature led to average increases of 0.7 °C for avg6-R, 0.6 °C for avg3-R, and 0.3 °C for avg1-R for a full dam steady-state.
- The influence of water temperatures at different water levels varies with time; this influence is governed by the water level. If the dam remains high, the water-temperature 60-day moving averages of the dam peak at different times. If the water level was low, avg1-R and avg3-R peaked at almost the same time while avg6-R peaked early. The amplitude of the cumulative water temperatures also varies, with avg3-R having the highest, then avg1-R, and lastly avg6-R.
- High water levels lead to low variances in the water temperatures, and low water levels lead to high variances in the water temperature. This is with exception to avg6-R, where there is no relationship between water level and avg6-R 60-day moving variance. These variances show the influence of different water temperatures with respect to time. Low variances mean higher influence, and high variances shows low influence.
- High water levels lead to a decrease in water-temperature moving average, and low water levels lead to an increase in water-temperature moving averages. Therefore, it is very important that extreme events like droughts and extended droughts as well as extended wet seasons are studied; this is incorporated with finite element analysis.

Author Contributions: P.M.: funding acquisition, resources, and supervision; T.T.: writing, software, formal analysis, and conceptualization; M.K.; supervision, writing review and editing, and resources. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fluixá-Sanmartín, J.; Altarejos-García, L.; Morales-Torres, A.; Escuder-Bueno, I. Review article: Climate change impacts on dam safety. *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 2471–2488. [CrossRef]
2. ICOLD. *Bulletin on Risk Assessment in Dam Safety Management*; International Commission on Large Dams: Paris, France, 2003.
3. ICOLD. *Risk Assessment in Dam Safety Management, a Reconnaissance of Benefits, Methods and Current Applications*; International Commission on Large Dams: Paris, France, 2005; Volume 30.
4. Bahls, V.; Holman, K. *Climate Change in Hydrologic Hazard Analyses: Friant Dam Pilot Study-Part I. I: Hydrometeorological Model Inputs*; Tech. Rep.; USA Bureau of Reclamation: Washington, DC, USA, 2014.

5. Bowles, D. Advances in the practice and use of portfolio risk assessment. In Proceedings of the ANCOLD Conference on Dams, New South Wales, Australia, 2000; Available online: https://www.ancold.org.au/?product=2000-advances-in-the-practice-and-use-of-portfolio-risk-assessment (accessed on 6 November 2020).
6. ICOLD. *Flood Evaluation and Dam Safety*; International Commission on Large Dams: Paris, France, 2018.
7. Pham, D.; Meding, J.; Kanjanabootra, S.; Luu, C. *Theoretical Framework for Dam Safety Risk Assessment for Vietnam*; RICS: London, UK, 2015.
8. Donnelly, R. International Waterpower and Dam Construction. Available online: https://www.waterpowermagazine.com/features/featuresafe-and-secure-risk-based-techniques-for-dam-safety-722/ (accessed on 4 September 2020).
9. National Research Council. *Informing Decisions in a Changing Climate*; National Academies Press: Washington, DC, USA, 2009.
10. IPCC. *Climate Change 2014: Impacts, Adaptation and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2014.
11. Escuder-Bueno, I.; Matheu, E.; Altaregos-Garcia, L.; Castillo-Rodrigues, J.T. *Risk Analysis, Dam Safety, Dam Security and Critical Infrastructure Management*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2011.
12. Atkins. *Impact of Climate Change on Dams & Reservoirs, Final Guidance Report FD2628*; Department of Environment, Food and Rural Affairs: London, UK, 2013.
13. Sheibani, F.; Ghaemian, M. Effects of Environmental Action on the thermal Stresses of Arch dam. *J. Eng. Mech.* **2006**, *132*, 532–544. [CrossRef]
14. Dandekar, A.; Basu, D.; Bressan, S. *Differentially Private Non-parametric Machine Learning as a Service*; Springer: New York, NY, USA, 2013.
15. Worms, J.; Touati, S. *Parametric and Non-Parametric Statistics for Program Perfomance Analysis and Comparison*; HAL: Paris, France, 2017.
16. Amberg, F. Interpretative models for concrete dam displacement. In Proceedings of the 23rd ICOLD Congress, Brasilia, Brazil, 25–29 May 2009.
17. Bishop, C. *Neural Networks for Pattern Recognition*; Oxford University Press: Oxford, UK, 1995.
18. Aggarwal, K.K.; Singh, Y.; Chandra, P.; Puri, M. Bayesian Regularization in a Neural Network Model to Estimate Lines of Code Using Function Points. *J. Comput. Sci.* **2005**, *1*, 505–509. [CrossRef]
19. Okut, H. *Bayesian Regularized Neural Networks for Small n Big p Data*; Intech: London, UK, 2016.
20. Heckerman, D.; Shortliffe, E. *Artificial Intelligence Med*; Springer: New York, NY, USA, 1992.
21. Zurada, J. *Introduction to Artificial Neural Syetem*; PWS: Boston, MA, USA, 1992.
22. Salazar, F.; Morán, R.; Toledo, M.Á.; Oñate, E. Data-Based Models for the Prediction of Dam Behaviour: A Review and Some Methodological Considerations. *Arch. Comput. Methods Eng.* **2017**, *24*, 1–21. [CrossRef]
23. Cehn, B.; Chang, M. Land forecasting using support vector machines: A study on EUNITE competition 2001. *IEEE Trans. Power Syst.* **2004**, *19*, 1821–1830.
24. Cheng, L.; Zheng, D. Two online dam safety monitoring models based on the process of extracting environmental effect. *Adv. Eng. Softw.* **2013**, *57*, 48–56. [CrossRef]
25. Hastie, T.; Tibshirani, R.; Friedman, J. *The Element of Statistical Learning Data Mining, Inference, and Prediction*, 2nd ed.; Springer: Berlin, Germany, 2009.
26. Meinshausen, N. Quantile regression forests. *J. Mach. Learn. Res.* **2006**, *7*, 983–999.
27. Held, I.; Pitman, A.; Planon, S.; Zhao, Z. *Evaluation of Climate Change Models*, 2012; in press.
28. Przyborski, P. Climate Change. In *The Scientific Basis*; Cambridge University Press: Cambridge, UK, 2005.
29. Weigel, A.P.; Knutti, R.; A Liniger, M.; Appenzeller, C. Risks of Model Weighting in Multimodel Climate Projections. *J. Clim.* **2010**, *23*, 4175–4191. [CrossRef]
30. Bellenger, H.; Guilyardi, E.; Leloup, J.; Lengaigne, M.; Vialard, J. ENSO representation in climate models: From CMIP3 to CMIP5. *Clim. Dyn.* **2014**, *42*, 1999–2018. [CrossRef]
31. Braconnot, P.; Harrison, S.P.; Kageyama, M.; Bartlein, P.J.; Masson-Delmotte, V.; Abe-Ouchi, A.; Ottobliesner, B.L.; Zhao, Y. Evaluation of climate models using palaeoclimatic data. *Nat. Clim. Chang.* **2012**, *2*, 417–424. [CrossRef]
32. Deser, C.; Phillips, A.; Bourdette, V.; Teng, H. Uncertainty in climate change projections: The role of internal variability. *Clim. Dyn.* **2012**, *38*, 527–546. [CrossRef]
33. Government of South Africa. *Representative Concentration Pathways*; Government of South Africa: Cape Town, South Africa, 2013.
34. Kozak, K.; Sare, L. Researching Climate Change: Trends in US Government Publications Distributed by the Government Printing Office: Washington, DC, USA, 2009. Available online: http://www.istl.org/09-fall/refereed4.html (accessed on 6 November 2020) US Library of Congress: Washington, DC, USA, 2009.
35. Van Vuuren, D.; Stehfest, E.; Elzen, M.G.J.D.; Kram, T.; Van Vliet, J.; Deetman, S.; Isaac, M.; Goldewijk, K.K.; Hof, A.; Beltran, A.M.; et al. RCP2.6: Exploring the possibility to keep global mean temperature increase below 2 °C. *Clim. Chang.* **2011**, *109*, 95–116. [CrossRef]
36. Mitchell, T.D. Pattern Scaling: An Examination of the Accuracy of the Technique for Describing Future Climates. *Clim. Chang.* **2003**, *60*, 217–242. [CrossRef]
37. NAST. Climate Change Impacts on the United States: The Potential Consequences of Climate Variability and Change, Report for the US Global Change Research Program. U.S. Global Climate Research Program. 2001. Available online: https://www.nrc.gov/docs/ML1006/ML100601201.pdf (accessed on 6 November 2020).

38. Miller, K.G.; Wright, J.D.; Browning, J.V.; Kulpecz, A.; A Kominz, M.; Naish, T.R.; Cramer, B.S.; Rosenthal, Y.; Peltier, W.R.; Sosdian, S. High tide of the warm Pliocene: Implications of global sea level for Antarctic deglaciation. *Geology* **2012**, *40*, 407–410. [CrossRef]

39. Zeebe, R.E.; Ridgwell, A.; Zachos, J.C. Anthropogenic carbon release rate unprecedented during the past 66 million years. *Nat. Geosci.* **2016**, *9*, 325–329. [CrossRef]

40. Xie, S.-P.; Deser, C.; Vecchi, G.A.; Collins, M.; Delworth, T.L.; Hall, A.; Hawkins, E.; Johnson, N.C.; Cassou, C.; Giannini, A.; et al. Towards predictive understanding of regional climate change. *Nat. Clim. Chang.* **2015**, *5*, 921–930. [CrossRef]

41. SANS, 204-2. Energy Efficiency in Buildings. SABS Standards Division: Pretoria, South Africa, 2011.

42. Moyo, P.; Oodthuizen, C. Sturctural Health Monitoring of Arch Dams Using Dynamic and Static Measurements; WRC Report No. 2025/1/13. 2013. Available online: http://www.wrc.org.za/wp-content/uploads/mdocs/2025-1-13.pdf (accessed on 6 November 2020).