Long Short Term Memory and Water Behaviours: A Lake Geneva Study

Malik Lechekhab
HEC Lausanne
MSc. in Finance
malik.lechekhab@unil.ch

Florence Hugard
HEC Lausanne
MSc. in Finance
florence.hugard@unil.ch

***Abstract*—Over the last decade, major hurricane and rainfall events devastated numerous regions of the world. Proven to be consequences of global warming [3], these unfortunate events have tremendous human and financial costs. In Switzerland, flooding and landslides caused CHF 170 million of damages in 2017 and about 66% of them were caused by thunderstorms and heavy precipitations [21]. Today there is a critical need to raise more awareness on how global warming could impact cities and populations. In the long term, the goal of this project is to provide predictions on the elevation and movements of oceans, rivers and lakes. This first study only focuses on lakes and more especially on Lake Geneva. The data used were daily water elevation data from January 1886 to April 2019 provided by the Swiss Federal Office for the Environment and have been applied to a 2-layer long short term memory (LSTM) network with a sliding window. Well-suited for time-series prediction, LSTM generally provides good results using less data than traditional water elevation prediction models. After having chosen the model's parameters through hyperparameter tuning, the model has been implemented on short-term, medium-term and long-term horizons and compared to a Naive model. Short-term and medium-term LSTM models do not seem to have a substantial predictive power when compared to Naive models. With long-term horizons, LSTM models perform better than Naive models showing higher predictive power for lakes water elevations. They however seem to underestimate flash water rise.**

***Index Terms*—Long short term memory, increasing sea levels, floods, neural networks, Lake Geneva**

I. Introduction

Every year climate change seriously impacts an estimated 325 million people worldwide [1]. Forming geological, biological and ecosystem alterations, climate change is leading to numerous environmental risks, such as rising sea levels, extreme weather and wildlife extinction or relocation [2]. These risks are impacting human lives in many different ways such as food supply, health, migration, drinking water and economic growth. One important risk that climate change has is rising sea levels, which in turn impacts hurricane and rainfall activities. As a consequence, the maximal potential energy that storms can release and the reach of rainfalls and storms increase drastically, causing more devastating hurricane and rainfall events. In 2018, Hurricane Florence provoked terrible casualties in the Carolinas with $24 million destruction costs and 52 direct and indirect fatalities. During the heaviest precipitating parts the hurricane rainfalls increased by over 50%. This increase is larger than thermodynamics expectations and would hence be due to human influence on climate change and warmer sea temperature [3]. When focusing on Switzerland, storms cause on average damages of CHF 307 million every year [1] [21]. Moreover, a Cornell University research estimated 2 billion people could become climate change refugees by 2100 due to rising sea levels [4]. Therefore, if flooding events could be forecasted in advance and with high accuracy, the risks and negative impacts associated could be mitigated. As flooding risks are expected to rise dramatically in the next decades, and more especially in the United States, Central Europe and North East and West Africa [5], there is a serious need for a predictive flooding model that could be adapted to any water surface. The goal of this project is to provide localities, institutions and individuals with such an open source flooding model only needing daily elevation data of the studied water surface. The first part of this project - that is described in this paper - focuses only on lakes' water elevation using Lake Geneva as a basis. In the long term, it would be interesting to extend the project to rivers and oceans. The data used for this paper were Lake Geneva daily water elevation from January 1 1886 to April 1 2019 provided by the Swiss Federal Office for the Environment (FOEN) and have been applied to a 2-layer long short term memory (LSTM) network with a sliding window. The LSTM deep learning method has been initiated by Juergen Schmidhuber and Sepp Hochreiter in the mid-90s and is a type of recurrent neural networks that can process any kind of data [15]. It is widely used by technology companies to develop new products ranging from speech recognition engines to automatic translation and music composition. Also well-suited for time-series prediction, LSTM generally provides good results using less data than traditional water elevation prediction models.

In response to water challenges, this paper will first provide background to the climate change issues and the different techniques that can be used to forecast water behaviours. Then it will discuss the methodology behind the LSTM as well as the choice of parameters used for the modelization.

[1] Inflation-adjusted average between 1972 and 2017

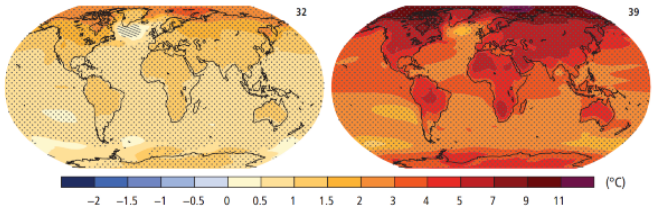
Finally, it will go over the results obtained using different sliding windows and delays and conclude with the further improvements that could be implemented in the future.

II. Literature Review

A. Climate change and increasing sea levels

According to IPCC (Intergovernmental Panel on Climate change), global sea level will rise by up to 60 cm by 2100 as a result of global warming and glaciers melting. However, as polar ice sheet mass is declining at an accelerated pace, the sea level rise (SLR) might even be 1 meter or more by 2100. Tide gauge data since the late nineteenth century shows that levels have risen by an average of 1.7 ± 0.3 mm per year since 1950. In the early 1990s, because of technical advancements, high-precision altimeter satellites began to measure SLR. The mean rate from 1993 to 2009 amounts to 3.3 ± 0.4 mm per year. [6]. This increase in sea levels is the result of temperature rising that has characterized the past decades. Compared to the late nineteenth century the Earth is about 0.8°C warmer today, meaningfully impacting the environment. If the emissions of greenhouse gases such as carbon dioxide are not brought under control, temperature rise is expected to be much larger in the future and is currently having severe consequences such as (1) thermal expansion of sea waters, caused by ocean heating; (2) melting mountain glaciers, adding significant amounts of water to oceans; (3) disintegrated ice sheets in Greenland and Antarctica, especially at their warmer peripheries. [6]

Fig. 1. Change in average surface temperature - 1986-2005 to 2081-2100



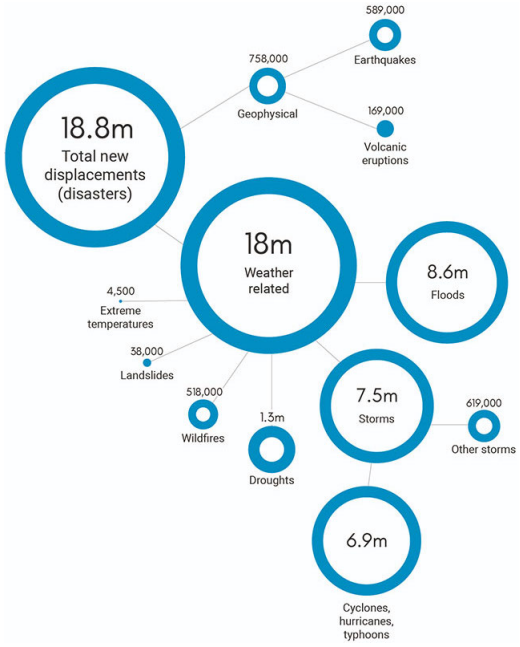
Source: Michael Oppenheimer, Adapting to Climate Change: Rising Sea Levels, Limiting risks. In: Social research Vol.82: No.2, 2015. Notes: 32 and 39 indicate numbers of models used to make each projection

B. Consequences in coastal areas

Heavy flooding and rainfall activities are amongst the most important impacts of SLR. If added to that sea levels keep increasing at an accelerated rate, flooded areas might even be larger than expected [9]. Except if mitigation and adaptation strategies are implemented by localities, coastal populations will experience a significant increase in the frequency of minor and major floodings over different SLR levels [10]. With a sea levels increase of 61 centimeters, major floodings are expected to occur multiple times a year and minor floodings more than 150 days a year [10]. Furthermore, Li et al. (2009) estimate that a one-meter sea level rise - plus a 10% surge intensification - would put 67 million people at risk. [7] Whereas, Bamber and Aspinall (2013) estimated this number to as high as 187 million people in this century. Hinkel et al. (2014) found that if the sea level rises with 1.2 meter, it can put up to 4.6% of the world global population in danger - i.e. around 345 million people.

Low-elevation coastal zones are typically very populated. In 2000, about 630 million inhabited these zones. These population might grow up to 1.4 billion by 2060. Today, most of the coastal areas with very high concentration of population are located in Asia [8]. As an order of magnitude new internal displacements caused by disasters represented 61% of new displacements in 2017 compared to 39% caused by conflicts. This represented 18.8 million displacements and 18 million were weather-related (see Figure 2) [19]

Fig. 2. Breakdown of new internal displacements due to natural disasters in 2017



Source: Internal Displacement Monitoring Center

C. Consequences of floods

Because of global warming, rainfall is shifting patterns, creating heavy rains and floods more frequently in many areas of the world. In addition, human alterations of land, such as engineering of rivers, destruction of natural protective systems, and increased construction in areas of low-lying ground adjacent to a river, are contributing to higher risks of experiencing destructive floods. Considered amongst the costliest natural disasters, river floods have many consequences such as property destruction, fatalities, contamination of surface water or inundation of dry land. Those can in turn lead to poverty, changes in biodiversity or waterborne diseases such as cholera [12].

D. Machine Learning: A Tool for Water Level Forecasting

Governments are more and more under pressure to establish reliable and accurate mapping of flood-risk areas and further develop prevention and protection strategies. Nonetheless, the dynamic nature of weather conditions makes the flooding events' time and location fundamentally complex to predict. Although physically-based methods have long been used to predict hydrological and climatic events and showed robust results for flooding scenario predictions, they generally use complex hydrogeomorphological datasets which require intensive computation and prevent short term forecasts. Data-driven statistical methods such as autoregressive moving average or multiple linear regression have also been at the center of flood modelization. They assimilate climate and hydrometeorological measures and are generally more computation cost-efficient compared to physically-based methods. However, some of these methods have been reported to be unsuitable for short term forecasts and lacking of accuracy. The previously-mentioned drawbacks of physically-based and statistical models are promoting the implementation of advanced machine learning (ML) techniques to forecast hydrometeorological events. ML methods have several advantages compared to them, the most important being that ML can "numerically formulate the flood nonlinearity, solely based on historical data without requiring knowledge about the underlying physical processes." [11]. Additionally, Tayfur et al. (2018) study on flood hydrograph predictions showed that ML methods make good predictions using less data and can therefore be adopted in poorly gauged measurement stations, often present in developing countries [13]. Deep learning methods can even use images and videos to predict water movements. Isikdogan, Bovik & Passalacqua (2017) proposed a convolutional neural network model learning characteristics of water surface. Using satellite images as inputs, the final trained model is able to separate water from any kind of other visible elements such as snow, clouds, land and ice [14].

E. Conclusion

Sea level has raised significantly in the past decade and is expected to rise in accelerated pace in the future. Heavy flooding and rainfall activities are amongst the most important impacts of SLR. If added to that sea levels keep increasing at an accelerated rate, flooded areas might even be larger than expected, putting tens of millions of people at risk. As floods are considered amongst the costliest natural disasters, it is necessary to find a reliable and accurate mapping of flood-risk areas predicting water movements of any kind of water surface such as oceans, rivers or lakes. Several tools exist: physically-based and statistical methods and machine learning methods. The first two show notable drawbacks that ML models counteract. Hence machine learning and more especially deep learning methods seem to be the most appropriate ones in order to forecast water movements of different kinds.

III. Methodology - Long Short Term Memory

A. Definition and Use

Long short-term memory (LSTM) is a type of recurrent neural networks (RNN) architecture initiated by the German researchers Juergen Schmidhuber and Sepp Hochreiter in the mid-90s [15]. LSTM networks can process any kind of data points such as images, speech, video or text data and is therefore used in various fields. In 2002, the first blues melodies are created using LSTM. The network successfully learnt a type of blues music and composed a proper musical structure of this style [16]. Aside from musical composition, this type of neural network is also well-suited for time-series prediction, handwriting recognition or even code generation. Nowadays major technology companies such as Apple, Google or Amazon are putting LSTM neural networks at the center of their products. For instance, Google recently created the Google's Neural Machine Translation system consisting of a deep LSTM network and is now using this technology for speech recognition for its smart assistant Allo and Google Translate [17].

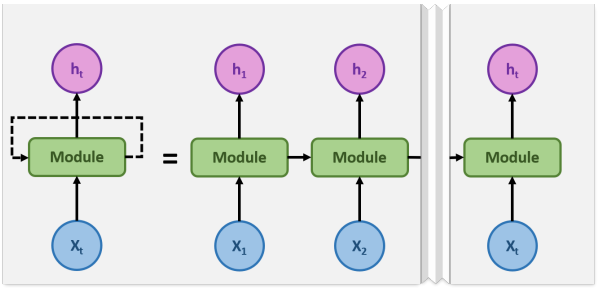
B. Why LSTM instead of traditional RNN?

The reason of the enthusiasm around LSTM is its ability to predict anything while dealing with the problems encountered in vanilla RNN - i.e. vanishing and exploding gradient [18]. In such problems, the gradient of the loss function approaches to zero - for vanishing gradient - or tends to infinity - for exploding gradient -, preventing the weight adjustment and making the network hard to train. As LSTM remembers information over a long period of time and assigns their importance through weights, gradient descent-based problems are avoided. Just as humans, LSTM networks do not start their thinking every second and connect previous information to current tasks. For instance, if one is trying to predict the last word in the sentence "machine learning is super cool", one does not need additional context as it is obvious that the coming word will be "cool". Most RNN can predict using past information but in some cases they might need more context - i.e. more sentences or words. In theory, it would enable them to learn more information but in practice it might not work. In those cases, LSTM become the leading choice. For these reasons, LSTM neural networks seem to be the appropriate method to predict Lake Geneva water elevation.

C. Framework

The structure of LSTM networks has the same base as any RNN, which are "dynamical systems with temporal state representations" [20], meaning that they consists of loops allowing information to persist in the system (see Figure 3).

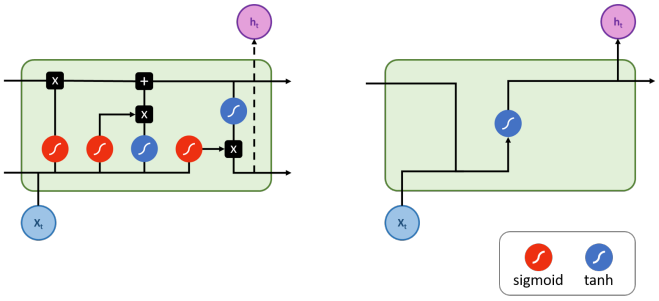
Fig. 3. Structure of a Recurrent Neural Network



Source: Fluss Creative Team

As mentioned above, the difference between LSTM and RNN comes from long-term dependencies, handled in LSTM but not in RNN - in practice. This difference is embodied in the structure of the LSTM itself. Indeed, RNN generally have a very simple structure of one tanh layer, whereas LSTM networks' structure is much more complex (see Figure 4).

Fig. 4. Structures of a LSTM (left) and a traditional RNN (right)



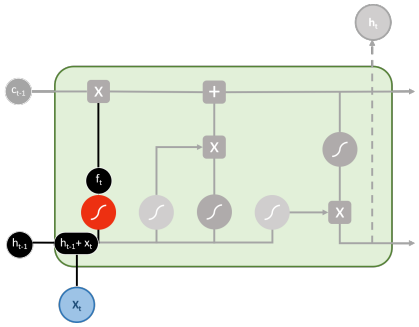
Source: Fluss Creative Team

LSTM are composed of a repeating module, which is itself composed of a cell - acting as the memory card of the module - and of three gates: the forget gate f_t, the input gate i_t and the output gate o_t [18]. Walking through the network, the first step is to decide which information are going to be thrown away from the cell. It is done by the sigmoid forget gate layer applied on the output of the previous module h_{t-1} and the input x_t (see Equation 2 and Figure 5). Composed of two parts, the second step determines which information are going to be stored in the cell. First, the sigmoid input gate layer decides which values are going to be updated (see Equation 3 and Figure 6). Secondly, a tanh layer creates new values $\hat{C}_t$ that could be added to the cell (see Equation 4). The combination of the two layers updates the cell's information. The new cell state becomes C_t (see Equation 1 and and Figure 7).

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \hat{C}_t \tag{1}$$

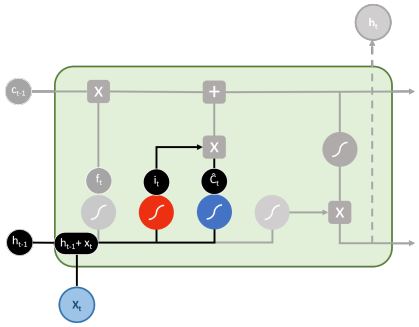
With f_t the sigmoid forget layer, C_{t-1} the cell state at $t-1$, i_t the input gate layer and $\hat{C}_t$ the the new values that could be added to the cell state. These are defined as followed.

Fig. 5. Step 1: Forget gate



Source: Fluss Creative Team

Fig. 6. Step 2: Input gate



Source: Fluss Creative Team

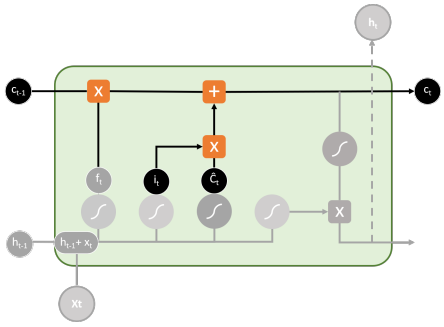
$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\hat{C}_t = \tanh(W_{\hat{C}} \cdot [h_{t-1}, x_t] + b_C) \tag{4}$$

With W_k the weights applied to the inputs h_{t-1} and x_t in the associated activation function and b_k the bias used compute element k with $k = \{f, i, \hat{C}\}$

Fig. 7. Step 3: Cell State update



Source: Fluss Creative Team

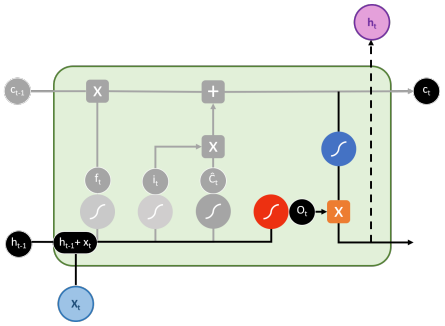
Finally, the module delivers the output h_t based on the sigmoid output layer o_t and the new cell state information

C_t put through a tanh activation function (see Equation 6 and Figure 8).

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{6}$$

Fig. 8. Step 4: Output gate



Source: Fluss Creative Team

In this study, the network used is composed of 2 LSTM layers and a linear Dense layer (see Figure 9). This last linear layer permits to obtain an output of size 1 - i.e. the water elevation y.

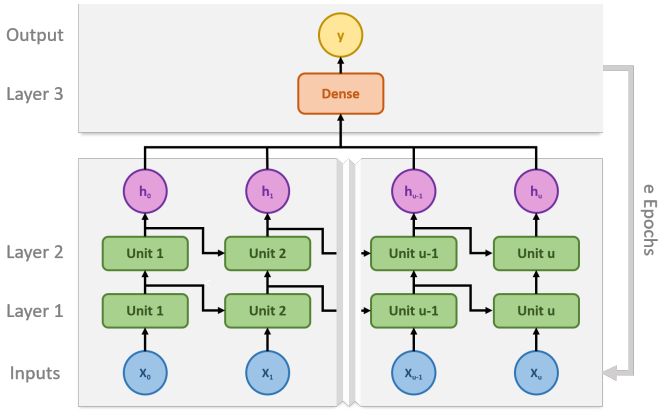
D. Parameter choice

The choice of parameters is a substantial component of neural networks modelization. The goal is to establish a prediction at $t + n$ of water elevation given k data available at time t. Therefore, a sensitivity analysis will be conducted in Section VI and will be based on a certain delay n and a sliding window of size k. However it is necessary to fix certain parameters such as the optimizer and the learning rate LR, the number of hidden layers of the model HL, the number of units of each hidden layer u, the number of epochs e and the batch size BS.

In neural networks, the learning rate is a decisive parameter used in gradient descents. It represents the step size and controls how slowly or quickly a neural network (NN) trains. In the proposed model, Adam optimizer has been selected as it showed huge speed performance and smaller losses compared to traditional optimizers such as AdaDelta or stochastic gradient descent (SGD) [22]. To this optimizer has been arbitrary associated the learning rate of 0.001 - which is also the default value in TensorFlow Keras [23].

In artificial neural networks, hidden layers are layers between the input and the output layers. For a single-layer network, the neuron takes in a set of weighted input and generate an output through a specified activation function. For the proposed model, a 2-layer LSTM network has been chosen. This structure showed more precise results than single-layer LSTM and, as the complexity of the dataset is fairly low, it did not seem necessary to add more layers.

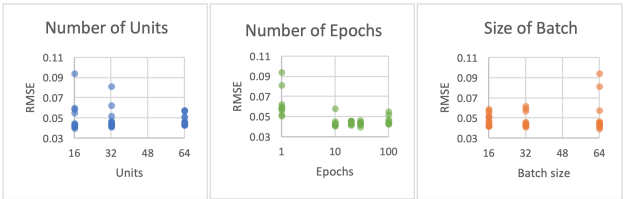
Fig. 9. LSTM network



Source: Fluss Creative Team

As shown on Figure 9, each layer is composed of an identical module repeated several times. Here the number of units u refers to the number of times it is repeated. Once the entire dataset has been through the network, it has completed one epoch. When setting the number of epochs e, one sets how many time the network should be trained with the data. Going through the network, the dataset can also be splitted in batches of the same size BS. This last parameter will define the number of samples to work through before updating the internal parameters of the model. In this study, the number of units per layer, the number of epochs and the batch size have been determine using hyperparameter tuning. To determine the optimal value of each parameter, a first range of values has been established such that $u = \{16, 32, 64\}$, $e = \{1, 10, 20, 30, 100\}$ and $BS = \{16, 32, 64\}$. The unit and batch size ranges have be chosen following a binary system of base 2 whereas the epoch range has arbitrary been defined after different tries [23]. The model has then been trained according to each possible set of parameters and the learning rate and number of hidden layers specified above. For each scenario, the model has finally been tested given a delay of $n = 7$ and a sliding window of $k = 30$ and the associated root mean square error (RMSE) has been computed.

Fig. 10. Root Mean Square Errors depending on different parameters' values

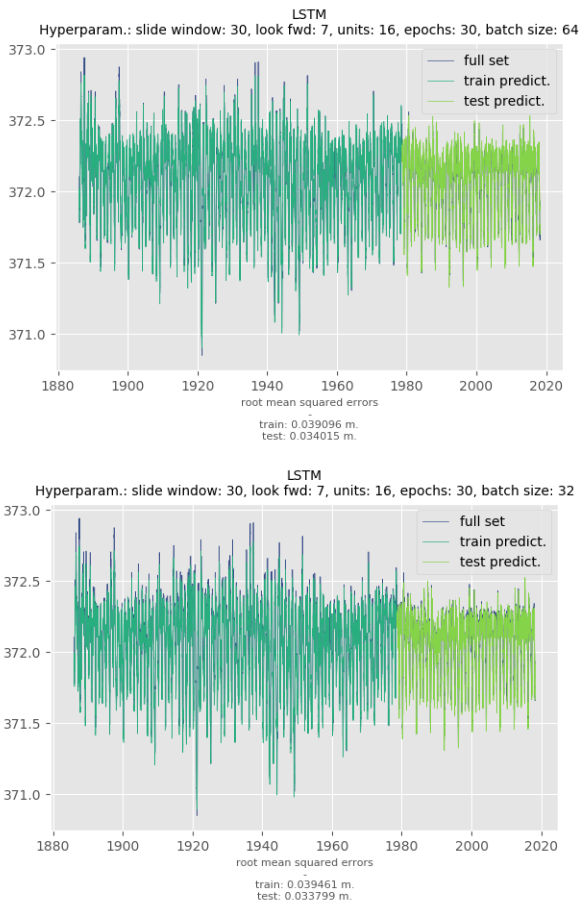


Source: Fluss & FOEN

At first, the parameters chosen were the ones of the model with the lowest RMSE and were as followed: $u = 16$, $e = 30$ and $BS = 64$. However when plotting RMSEs to the different

values of each parameter the batch size trend seemed to be different, showing a flatter RMSE distribution for $BS = 64$ compared to the other values $BS = \{16, 32\}$ (see Figure 10) [23]. To verify the result obtained, the mean of the RMSE for each parameter value has been calculated. When minimizing this mean for each parameter, the following optimal set is obtained: $u = 16$, $e = 30$ and $BS = 32$, better matching expectations (see Excel file `hyperparameter_tuning.xlsx`) [23]. To validate the most accurate option, both resulting models have been applied to the full-dataset training and testing sets of size 70% and 30% respectively. For the first model the training and testing sets were showing RMSEs of 0.040818 and 0.035040 respectively whereas they were showing RMSEs of 0.039444 and 0.033745 for the second (see Figure 11). Consequently, the chosen parameters were the one of the second model: $u = 16$, $e = 30$ and $BS = 32$.

Fig. 11. Mean Square Error of the different scenarios



Source: Fluss & FOEN

IV. Dataset Description

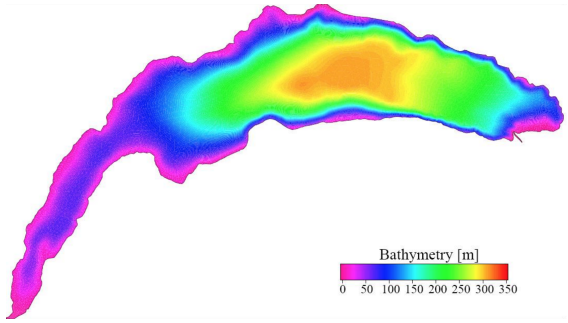
A. Lake Geneva

Lake Geneva has been chosen as the first study of water surface because it gathers several advantages: proximity and data accessibility. Indeed, as University of Lausanne students, it is much easier to consult one-site experts of Lake Geneva if needed. Data extraction has been facilitate by the Swiss Federal Office for the Environment (FOEN) which promptly provided the data.

Geography: Lake Geneva or Lac Léman in French is a glacial lake situated between Switzerland and France at the North of the Alps. Considered as one of the largest lakes in Europe, Lake Geneva covers an area of 7975 squared meters, has a maximal length of 75 kilometers and a maximal width of 14 kilometers. It also have an average depth of 154 meters and a maximum depth of 310 meters (see Figure 12) [25]. Lake Geneva has several affluents coming from different rivers in France and Switzerland, the most important being the Rhone River which takes its source in the Rhone Glacier.

Climate: During the year, the daily average of water temperatures of the lake vary between 1.5°C in January and 20.2°C in July [26]. Started in 1957, Lake Geneva water temperature tracking revealed an increase in deep water and average surface water temperatures. Supposedly due to climate change, these increases amount to +1.1°C between 1963 and 2016 for deep water and to +2°C between 1970 and 2016 for surface water, having impacts on the lake ecosystems [27].

Fig. 12. Bathymetric Map of Lake Geneva



Source: Cartes bathymétriques du Léman [25]

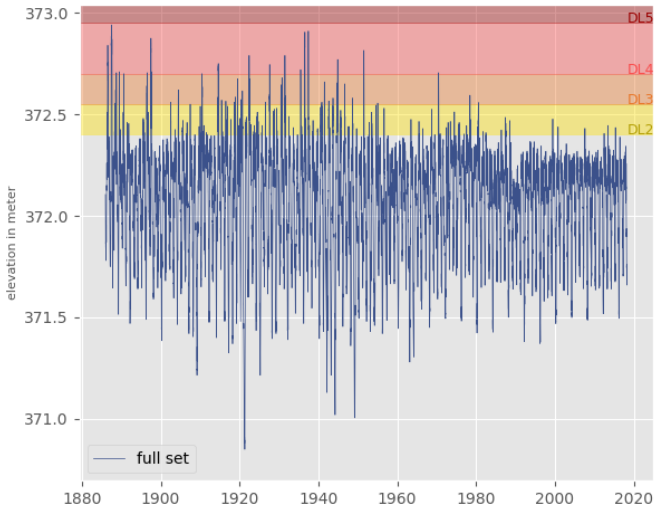
B. Measurements

The Federal Office for the Environment: The Swiss Federal Office for the Environment (FOEN) ensures the sustainable use of natural resources such as soil, water, air, quietness and forests in Switzerland. Thanks to their 349 stations spreaded across the country, the FOEN can record temperatures, water elevations and other hydrologic metrics [28]. Lake Geneva has 2 stations, one in St-Prex and one in Geneva - where

the data used come from. Additionally to the metrics, the FOEN has also established danger levels scaling from 1 to 5 corresponding for Lake Geneva to levels lower than 372.4 meters to more than 372.95 meters (see Figure 13).

Water measurement methods: In Switzerland, water levels are established through periodical measurements realized in permanent measuring stations. To do so, both non-recording and recording instruments can be applied. As of today, the trend goes toward electronic data recording and two main methods are used by the FOEN: radar systems and gauges with floats [29]. Non-interventional radar systems determine water levels without making direct contact with water and therefore are insensitive to branches and other rubbles. Gauges with floats measure the water elevation based on the displacement of the floats. For example, if the water level drops, the float will fall and the tape - connecting the float and the recording instrument - will be pulled by the weight of the float and send a signal to the instrument. Although they are fairly expensive to install and maintain, and must be protected from the tree branches, bed load and other rubbles, gauge with floats is the most widely used method.

Fig. 13. Time series of water elevation of Lake Geneva from 1886 to 2019



Source: Fluss & FOEN

C. Data

The dataset used in this project is a time serie of water elevation measurements of Lake Geneva. The data are expressed in meter with millimetric precision ranging from January 2 1886 to April 1 2019, which corresponds to 133 years of daily data and 48'666 observations. The data were kindly delivered on request by the FOEN in the form of .asc files containing daily average of elevation measurements. The files were then converted to .csv to ensure easy reading by the Python libraries.

Descriptive Statistics: Over the last 133 years, the average level of Lake Geneva was of 372.06 meters and this value has not significantly changed over the last 10 years. However, it appears that the variance was relatively higher before 1980 with a standard deviation of 0.30 meters versus 0.21 meters after 1980. The most likely explanation for this result is that the measurement tools have been updated at this time, allowing for more precise measurements. A first autocorrelation verification shows that the data are autocorrelated, which is expected for a process as periodic as water elevation.

V. Implementation

A. Importation and cleaning

The first step is to import the data and split them in two different files. The first file called `elevation_PRES.csv` contains the training and testing sets' data that will be use for the training. The second one called `elevation_FUT.csv` contains the out-of-sample data that will be use in Section VI. Then the data need to be formatted to be usable on Python. The formatting includes to (1) drop the first column corresponding to the postal code of the measurement station (2) reformat the dates and (3) setting them as index column.

B. Normalization

The dataset is normalized in order to ease the use of sigmoid function and for the sake of comparison. This step transforms the features by scaling each value to a range between the dataset minima and maxima. The transformation is an alternative to zero mean and unit variance scaling.

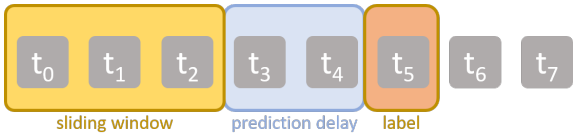
C. Splitting

The data are splitted in two different sets: training set and testing set. The respective proportions are 70% and 30%. The 70/30 rule is commonly used for training and testing models.

D. Creation of sliding window matrix

The function `create_dataset()` is then call to create a matrix of shifted inputs and corresponding vector of observations for a given sliding window and forecast delay. For instance, a sliding window of 3 means that the model studies 3 observations preceding the value to predict. Then the prediction delay define how far in time the prediction should be. If the delay is 2, it means that the model will studies the observations at t_0, t_1 and t_2 in order to predict the value in t_4.

Fig. 14. Sliding window, delay and prediction



Source: Fluss Creative Team

E. Model setting and compilation

Once the data are ready, the model is defined - here as sequential. It is composed of two LSTM layers of n units and a dense layer, which will generate the final outputs. Secondly, the model is compiled with the mean squared error as loss function and the Adam optimizer. Finally the model is fitted with the training set and testing set - here also defined as the validation dataset. The number of epochs and the batch size are defined at this point. Then the mean root square error of the training predictions and of testing predictions are computed.

F. Presentation of the results

To assess the performance of the model, it is applied to out-of-sample data. All the formatting described above are applied again before running the model. In order to present the results and be correctly interpreted, the values are denormalized and shown through three main plots. The first one shows an overview of the sample on which the model is trained with the different water elevation danger levels (see Figure 13). The second plot shows the results of the training and testing processes (see Figure 11). Finally the third plot is composed of three subplots: (1) one showing the results of the LSTM model, the results of the Naive prediction model and the values observed; (2) one showing the history of the losses of the training and testing sets both converging and (3) one comparing the RMSEs of the LSTM model and Naive model.

VI. Results

The goal of this study is to assess lakes water behaviours taking into account different time horizons. Initialized with the hyperparameters defined in Section III-D, the model generates divergent results depending on the time horizon that is considered. The long-term forecast seems to be more accurate than the short-term and medium-term forecasts. Note that the results observed come from an out-of-sample dataset containing one year of data from April 1 2018 to April 1 2019. The model has been trained and tested on data from January 1 1886 to March 31 2018, with a 70/30 split. Additionally to the LSTM model, a Naive model has been implemented for the sake of comparison. In this last model, the prediction of Lake Geneva water elevation in n days should be the water elevation of today. If the LSTM model shows better accuracy than this very simplistic model it means that it has a higher power of prediction.

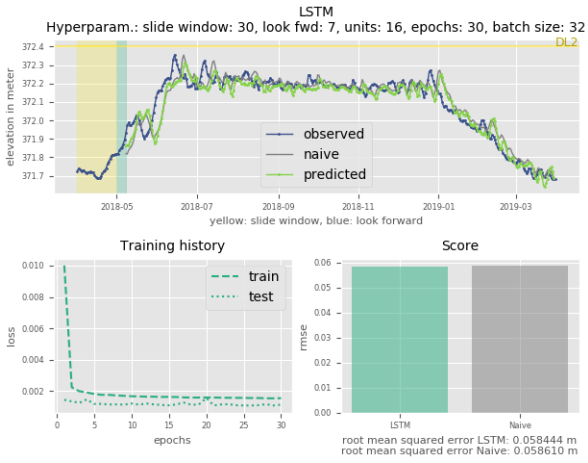
Fig. 15. Short-term forecast with $k = 7$ and $n = 3$ on one-year data

Source: Fluss & FOEN

A. Short-term forecast

Short-term forecasts represent predictions using a sliding window of one week - i.e. 7 days - and a delay of 3 days. Figure 15 shows the resulting forecast as well as the associated model loss over the training and testing sets, and the RMSE of the LSTM model and the Naive model for the given hyperparameters. Even if the loss graph converges, it seems that the Naive model has a higher prediction power than the LSTM model, with a RMSE of 0.038123 and 0.044655 respectively.

Fig. 16. Medium-term forecast with $k = 30$ and $n = 7$ on one-year data



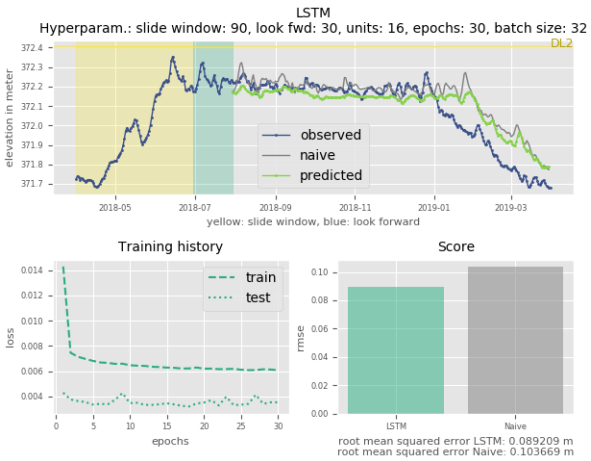
Source: Fluss & FOEN

B. Medium-term forecast

With a sliding window of one month - i.e. 30 days - and a delay of 7 days, the medium-term forecast shows

more encouraging results (see Figure 16). As before, the loss function converges for both the training and testing sets. However, the RMSE of the LSTM model is slightly lower than the one of the Naive model, with a value of 0.058444 and 0.058610 respectively. This difference is of order 10^{-4} meter which is particularly low. This fairly low result might be surprising as the hyperparameter tuning has been conducted with similar sliding window and delay values.

Fig. 17. Long-term forecast with $k = 90$ and $n = 30$ on one-year data

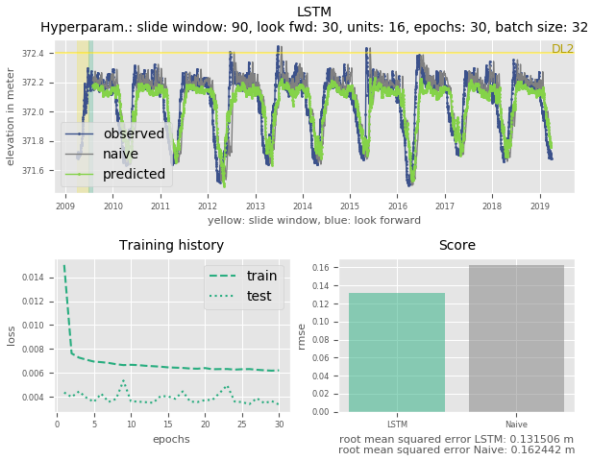


Fluss & FOEN

C. *Long-term forecast*

Long-term forecasts are using a sliding window of three months - i.e. 90 days - and a delay of 30 days. Figure 17 shows the convergence of the loss function for the training and the testing sets and very interesting results regarding the RMSE of the LSTM and Naive models. Indeed, LSTM presents a RMSE value of 0.089209 whereas the Naive model shows a RMSE of 0.103669 - a -13.95% difference - meaning that the LSTM model has a higher prediction power for water elevations. To verify this results, a similar model has been implemented, changing only the size of the out-of-sample dataset and consequently the size of the training and testing sets - always keeping a 70/30 split. The forecast now carried out on 10 years data from April 1 2009 to April 1 2019 has even more striking properties. As before, the LSTM model has a lower RMSE of 0.131506 compared to the Naive model which has a 0.162442 RMSE value, representing a -19.04% difference. However, the model seems to underestimate flash water rise. Over the last 10 years, the model did not predict correctly important peaks of water level that crossed the danger level 2 (see Figure 18). Although this LSTM should not be recommended for sudden water rise, its applications in long term prediction should be further explored.

Fig. 18. Long-term forecast with $k = 90$ and $n = 30$ on ten-years data



Fluss & FOEN

VII. Conclusion

Every year climate change seriously impacts an estimated 325 million people worldwide [1]. One important risk induced by climate change is rising sea levels, which in turn impacts flooding and rainfall activities. As flooding risks are expected to rise dramatically in the next decades and more especially in the United States, Central Europe and North East and West Africa [5], there is a serious need for a predictive flooding model that could be adapted to any water surface. Using LSTM neural networks, it has been possible to predict Lake Geneva water behaviors for long-term horizons with better accuracy than Naive model. However, the model seems to underestimate flash water rise. Additionally short-term and medium-term LSTM models do not seem to have a substantial predictive power when compared to Naive models. Hence, LSTM applications in long term prediction should be further explored using for instance satellite images. Adding another dimension - e.g. the water surface - could provide more insights on lakes water behaviours. An additional element that could be explored is the choice of parameters for long-term models. In this paper, the hyperparameter tuning has been implemented on a medium-term framework. With new hyperparameters, one could expect better performance of long-term LSTM models.

REFERENCES

[1] The anatomy of a silent crisis, (2009) Human impact report climate change, Global humanitarian forum Geneva

[2] European Commission website, (18 May, 2009) Retreived from: https://ec.europa.eu/clima/change/consequences_en

[3] Reed, K. A., Stansfield, A. M., Wehner M. F., Zarzycki, C. M.(2018), The human influence on Hurricane Florence

[4] Geisler C., Currens B., (2017), Impediments to inland resettlement under conditions of accelerated sea level rise. Land Use Policy, DOI: 10.1016/j.landusepol.2017.03.029

[5] Willner, S., N., Leverman, A., Zhao, F. Frieler, K., (2018) Adaptation required to preserve future high-end river flood risk at present, Science Advances

[6] Michael Oppenheimer, Adapting to Climate Change: Rising Sea Levels, Limiting risks. In: Social research Vol.82: No.2, 2015

[7] Li, X., Rowley, R.J., Kostelnick, J.C., Braaten, D., Meisel, J., Hulbutta, K., (2009). GISanalysis of global impacts from sea level rise. Photogram. Eng. Remote Sensing75, 807–818.

[8] Geislera, Ben Currens (2017) Impediments to inland resettlement under conditions of accelerated sea level rise, ScienceDaily. Retrieved on May 17, 2019 from www.sciencedaily.com/releases/2017/06/170626105746.html

[9] Mimura, N. (2013), Sea-level rise caused by climate change and its implications for society. Retrieved on May 25 from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3758961/

[10] Ghanbari, M., Arabi, M., Obeysekera, J. & Sweet, W. (2019), A Coherent Statistical Model for Coastal Flood Frequency Analysis Under Nonstationary Sea Level Conditions. Retrieved on May 25 from https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018EF001089

[11] Mosavi, A., Ozturk, P. & Chau, KW. (2018), Flood Prediction Using Machine Learning Models: Literature Review. Retrieved on May 25 from https://www.researchgate.net/publication/328562202_Flood_Prediction_Using_Machine_Learning_Models_Literature_Review.

[12] Bariweni, P.A., Tawari, C.C., & Abowei, J.F.N. (2012), Some Environmental Effects of Flooding in the Niger Delta Region of Nigeria. Retrieved from http://maxwellsci.com/print/ijfas/v1-35-46.pdf.

[13] Tayfur, G., Singh, V.P., Moramarco, T. & Barbetta, S. (2018), Flood Hydrograph Prediction Using Machine Learning Methods. Retrieved from https://www.researchgate.net/publication/326595868_Flood_Hydrograph_Prediction_Using_Machine_Learning_Methods.

[14] Isikdogan, F. & Bovik, A.C. (2017), Surface Water Mapping by Deep Learning. IEEE Journal. Retrieved on May 25 from http://www.isikdogan.com/files/isikdogan2017_deepwatermap.pdf.

[15] Hochreiter, S. & Schmidhuber, J. (1997), Long Short Term Memory. Retrieved from https://www.researchgate.net/publication/13853244_Long_Short-term_Memory.

[16] Eck, D., Schmidhuber, J. (2002), A First Look at Music Composition using LSTM Recurrent Neural Networks. Retrieved on May 25 from http://people.idsia.ch/~juergen/blues/IDSIA-07-02.pdf

[17] Wu, Y., Schuster, M., Chen, Z., Le, C.V., Norouzi, M. et al. (2016). Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. Retrieved from https://arxiv.org/pdf/1609.08144.pdf

[18] Olah, C., (2015), Understanding LSTM Networks, http://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[19] Internal Displacement Monitoring Centre. Global Report on Internal Displacement (2018). Retrieved from http://www.internal-displacement.org/global-report/grid2018/downloads/2018-GRID.pdf

[20] Du, K.-L., Swamy, M.N. (2013), Neural Networks and Statistical Learning. Retrieved from https://www.researchgate.net/publication/264912450_Neural_Networks_and_Statistical_Learning

[21] WSL Institute for Snow and Avalanche Research SLF (2008), Substantial flood and landslide damage in Switzerland 2017: Eight deaths and extensive material damage. Retrieved from https://www.slf.ch/en/news/2018/03/substantial-flood-and-landslide-damage-in-switzerland-2017-eight-deaths-and-extensive-material-damage.html

[22] Kingma, D.P. & Ba, J.L. (2015). Adam: A Method for Stochastic Optimization. Retrieved from https://arxiv.org/pdf/1412.6980.pdf

[23] Keras, Usage of optimizers. Retrieved from https://keras.io/optimizers/

[24] Keskar, N.S, Mudigere, D., Nocedal, J., Smelyanskiy, M., Tak, P. & Tang, P. (2015). On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima. Retrieved from https://arxiv.org/abs/1609.04836

[25] Loizeau, J.L. Cartes bathymétriques du Léman. Retrieved from https://www.unige.ch/forel/files/7614/1691/5984/Cartes_bathymetriques_disponibles_2a.pdf

[26] Federal Office of Meteorology and Climatology (2010),Climate Normals Geneva-Cointrin. Retrieved from https://www.meteoswiss.admin.ch/product/output/climate-data/climate-diagrams-normal-values-station-processing/GVE/climsheet_GVE_np8110_e.pdf

[27] Commission internationale pour la protection des eaux du Léman (2017). La Lettre du Léman, Juin 2017. Retrieved from https://www.cipel.org/wp-content/uploads/2017/06/LL54_FR.pdf

[28] Federal Office for the Environment FOEN - Hydrological data and forecasts. Stations and data. Retrieved from https://www.hydrodaten.admin.ch/en/stations-and-data.html

[29] Spreafico M. and Weingartner R. (2005). The Hydrology of Switzerland, Reports of the FOWG, Water Series. Retrieved from https://www.bafu.admin.ch/bafu/en/home/topics/water/state/water--monitoring-networks/basic-monitoring-network--water-levels-and-discharge-in-surface-.html