Accurate river level predictions using a Wavenet-like model

Shannon Doyle
University of Amsterdam
sjdoyle46@gmail.com

Anastasia Borovykh
Imperial College London
a.borovykh@imperial.ac.uk

Abstract

The effects of climate change on river levels are noticeable through a higher occurrence of floods with disastrous social and economic impacts. As such, river level forecasting is essential in flood mitigation, infrastructure management and secure shipping. Historical records of river levels and influencing factors such as rainfall or soil conditions are used for predicting future river levels. The current state-of-the-art time-series prediction model is the LSTM network, a recurrent neural network. In this work we study the efficiency of convolutional models, and specifically the WaveNet model in forecasting one-day ahead river levels. We show that the additional benefit of the WaveNet model is the computational ease with which other input features can be included in the predictions of river stage and river flow. The conditional WaveNet models outperformed conditional LSTM models for river level prediction by capturing short-term, non-linear dependencies between input data. Furthermore, the Wavenet model offers a faster computation time, stable results and more possibilities for fine-tuning.

Introduction

Due to climate change, there is an increasing occurrence of floods and droughts affecting river levels which poses a danger to human life, infrastructure and results in significant economic and social losses. Accordingly, accurate river level forecasts are needed for early warning systems of floods and flood mitigation, as well as the management of environment, infrastructure and shipping [11, 2]. River levels like the stage, the water height, and the flow rate are frequently measured along the river and can be estimated in forecasts. However, it is difficult to produce accurate river level forecasts because river levels are the result of complex, dynamic processes of the hydrological cycle. They are affected by many factors that determine water runoff into rivers such as topography, soil characteristics, land usage or snowmelt. Therefore river levels display large variations over time with a strong noise component, seasonality and a high degree of non-linearity. These characteristics pose challenges and uncertainty in parameter estimates to forecasting models [17].
Classical approaches for river level forecasting make use of historical time-series or river level related variables. Physically-based methods are hydrodynamic models. They attempt to model the mathematical process underlying the hydrological cycle using physical laws [15]. Hydrodynamic models require large amounts of input data, that can be difficult to obtain (i.e. topography or river geometry), need a long time to run and are prone to errors [19, 6]. A classical data-driven alternative are the auto-regressive (AR) models [10, 4]. However, they fail to capture the *non-linear* dynamics in river levels and are therefore outperformed by neural networks in river level forecasting [13].
Feedforward fully-connected neural networks can only access a small number of past time points for their predictions. They have been shown to struggle with generalisability on the test set and underprediction of values, because of the low occurrence of extreme events [1, 10, 7]. These models are also not able to encode temporal or sequential patterns of a time-series. The current gold-standard for time-series predictions is the LSTM, a recurrent neural network, which is specialised for this

Tackling Climate Change with Machine Learning workshop at NeurIPS 2020.

task; LSTM models have produced impressive results for river level forecasting, outperforming other models in forecasting one day ahead flow rates based on historical time-series of flow and catchment rainfall [11].
One of the challenges of RNNs and LSTMs is their training process which suffers from vanishing gradients and hampers learning long-time dependencies. Additionally they are not built to separately process conditional inputs. As an alternative to RNNs in time-series forecasting, convolutional neural networks (CNNs) have been proposed. While CNNs are known for their application to image classification, they can be applied to sequence learning as well [9, 5]. CNNs can make use of parallelisation and weight sharing among filters, resulting in enhanced computational efficiency, easier training and their ability to process long sequences. A specific CNN called WaveNet was developed for audio generation purposes [14]. The Wavenet model's architecture (see Figure 1) can access a broad history of data through dilated causal convolutions. It is able to identify short-term and long-term dependencies through its specialised residual block architecture and parameterised skip connections. Importantly, the Wavenet allows to handle additional inputs to condition its output. Because of the Wavenet's demonstrated ability to learn dependencies in the data at different scales in audio signals, it suggested itself to be useful in various applications in time-series forecasting problems [16, 3, 18, 20].
In this work, we assess the applicability and performance of a WaveNet-like network for single-step river level time-series forecasting by benchmarking it against the LSTM network and conclude that WaveNet models can be the new state-of-the-art in river level forecasts.

Methods

The Wavenet model is a CNN with dilated causal convolutions and parameterised skip connections as well as a specialised residual block architecture inspired by the LSTM cell.
Figure 1 shows that the dilated causal convolutions enable the network to access a broad history of the data, referred to as a large receptive field. Outputs from all layers are sent to the last layer via skip connections to allow deeper networks to converge. Each neuron of the hidden layers in Figure 1 (L) represents one residual block. The residual block architecture, showin in 2 (R) is inspired by the LSTM cell, which includes a hyperbolic tangent (tanh) gate and a sigmoid (σ) filter. The purpose of the residual connection within the block is to aid in the convergence of deep networks. The final output of the network is obtained by adding up all skip connections and then passing this output through a convolutional layer with a ReLU activation function and a convolutional layer with a softmax activation function. The Wavenet model output is the prediction of a single time-series. The value of the WaveNet model lies in the ability to condition this prediction on other supporting time-series by processing these with additional convolutional layers and adding them to the input of the first layer of residual blocks.

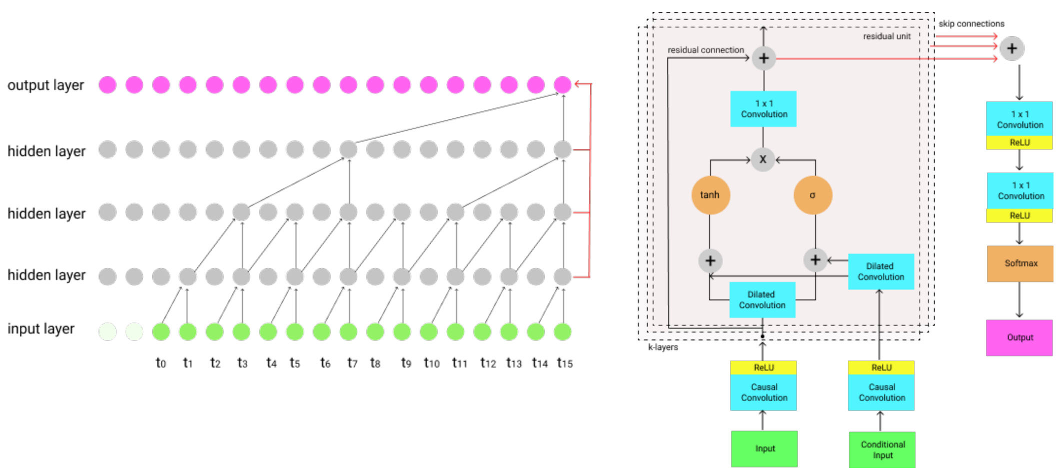


Figure 1: WaveNet architecture: (L) The causal dilated convolutions allow an output time step to receive information from a broader receptive field (green) with an increasing number of hidden layers. (R) Primary and conditional input is processed in the residual units and sent to output via skip connections. .

Data and Models

The models were trained for the station North Muskham on the river Trent. The stage data was obtained from the website River Levels UK (https://riverlevels.uk/rivers/rivertrent), and the river flow and daily catchment rainfall data were obtained from National River Flow Archive (https://nrfa.ceh.ac.uk/data/). The data included in this study (2012-2017) had no missing values. As a pre-processing step, all data types were min-max normalised. The conditional models predict daily river stage or daily river flow using rainfall and river flow or river stage, respectively. The hyper-parameter optimisation of models involved a variation of the number of layers and either the number of filters per layer (WaveNet) or the number of units per hidden layer (LSTM).

Performance Evaluation

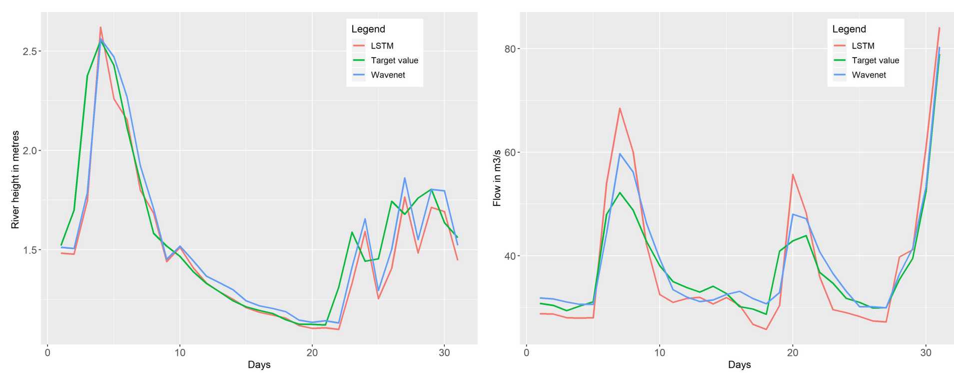
The RMSE was used as the loss function capturing the absolute error, while model performance was evaluated with the E-value and the U2-value. The E-value measures the goodness-of-fit and is computed like the well-known R^2 value used with statistical projections [12]. Theil's U2 value stems from econometrics and assesses the models' performance against a naive prediction. The U2 value is defined by,

$$U_2 = \frac{\sum_{t=1}^{n-1} \left(\frac{\hat{y}_{t+1} - y_{t+1}}{y_t} \right)^2}{\sum_{t=1}^{n-1} \left(\frac{y_{t+1} - y_t}{y_t} \right)^2}, \tag{1}$$

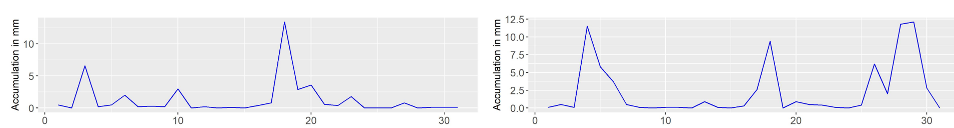
where y_t is the model input value at time t, as well as the naive prediction at time point $t+1$, y_{t+1} is the target value and $\hat{y}_{t+1}$ is the model prediction at time point $t+1$.

Results and Discussion

Table 1 and Figure 2 show the results of the WaveNet and the LSTM model.



(a) Unconditional Models: Stage predictions (b) Conditional Models: Flow predictions



(c) Rainfall during stage predictions (d) Rainfall during flow predictions

Figure 2: Predictions from WaveNet and LSTM models- unconditioned or conditioned with rainfall and river levels across two different periods of 30 days containing multiple peaks.

The unconditional models' predictions heavily relied on the target value of the previous time point, particularly when river levels rise. This is reflected by the models' high U2 values close to 1 (the value of a naive prediction), shown in table 1. The conditional models, however, generated realistic one-step-ahead predictions. The conditional models were able to *anticipate* peaks in the river levels

Table 1: Results best-performing Models

Model	Cond. Input	U2 mean	U2 Std	E Mean	E Std	RMSE Mean	RMSE Std
River Variable Stage							
WaveNet	-	0.921	0.004	**0.936**	7e-04	**0.152**	8e-04
LSTM	-	**0.888**	0.027	0.932	0.005	0.173	0.007
WaveNet	Rain, Flow	**0.525**	0.001	**0.967**	1e-04	**0.035**	1e-04
LSTM	Rain, Flow	0.596	0.022	0.962	0.003	0.038	0.002
River Variable Flow							
WaveNet	-	**0.918**	0.002	**0.930**	0	**19.727**	0.002
LSTM	-	1.084	0.225	0.928	0.007	20.285	0.995
WaveNet	Rain, Stage	**0.524**	0.006	**0.967**	0.001	**8.176**	0.147
LSTM	Rain, Stage	0.693	0.052	0.960	0.004	9.007	0.407

as is also reflected in lower U2 values. The WaveNet models significantly outperformed the LSTM models for the prediction of the stage and the flow in terms of U2-value, E-value and RMSE (Table 1). Figure 2 shows that the peaks in the primary variables often follow peaks in the rainfall data. The conditional models were able to anticipate peaks, using the information from this conditional input. The higher performance of the WaveNet models suggests that they were better able to extract the dependencies between the inputs than the LSTM. The WaveNet architecture offered possibilities to adjust the filter size at each of the convolutional gates in the residual units to extract information from the inputs optimally (see Figure 1). Such adjustments were not possible with the standard LSTM model. In some instances, the conditional models failed to predict peaks; in many of these instances, no peaks in the conditional input could be observed, suggesting that the model did not have access to sufficient information for predicting the peaks. This could likely be addressed by providing the model with additional (time-series) data of other conditioning factors. The hyper-parameter optimisation showed that the best models only required a small input window; for the WaveNet models, the receptive field was 8 days, while the training window of the LSTM was 28 days. This indicates that the critical information for the one-day ahead river level prediction lies in the recent past. The WaveNet likely did not benefit from its known ability to extract long-term dependencies when using a large receptive field, because the long-term dependencies were not sufficiently informative or representative (because of weight stationarity) to outweigh the introduction of noise by a larger receptive field. Further, table 1 shows that the WaveNet models' performance metrics had low standard deviations in comparison to those of the LSTM models. The WaveNet model produced such reliable and robust results, as it is able to process the whole sequence in one sample thanks to parallelisation and parameter sharing. The LSTM must process the input in smaller batches due to its vanishing gradients problem. Batching leads to inconsistencies in the output between separately trained models, because each batch has its own optimal solution of weights and the gradient update is a stochastic process [8]. These differences in architecture resulted in lower training times for the WaveNet models than for the LSTM models, showcasing the benefits of WaveNet models in practice, when fast and accurate river flow or stage predictions are needed.

Conclusion and Future outlook

We showed that the WaveNet architecture is a strong alternative to the LSTM for river level forecasting, as it achieves a lower prediction error by extracting short-term relationships between input data. The additional benefits of WaveNet include its reliability, possibilities for fine-tuning, and computational efficiency. The models' performance can likely be improved even further by adding suitable conditional inputs such as variables of the hydrological cycle or other influencing factors.

The model's reliable one-day head forecasts will allow densely populated areas located next to and downstream of major rivers to take emergency measures in the case of flooding. Furthermore, these short-term predictions are needed for controlling inland shipping and managing water supply. An advantage of the models is that they only require a small number of commonly available inputs. They can be trained and tested on the data of each location with the relevant and available conditional time-series. For usage of the WaveNet model in modern river level forecasting, the model should be validated on other stations and it can be tested for prediction of broader forecast horizons, which are needed for other areas of application, including various types of planning such as city planning or long-term flood mitigation.

Acknowledgements

Shannon would like to thank Emma Harvey and Bardo Bakker for supervision of this research performed for HAL24K Amsterdam in the Netherlands.

References

[1] Robert J Abrahart and Linda See. Neural network vs. arma modelling: constructing benchmark case studies of river flow prediction. pages 17–19, 1998.

[2] Stefano Alvisi, Gianluca Mascellani, Marco Franchini, and A Bardossy. Water level forecasting through fuzzy logic and artificial neural network approaches. 2006.

[3] Anastasia Borovykh, Sander Bohte, and Cornelis W Oosterlee. Conditional time series forecasting with convolutional neural networks. *arXiv preprint arXiv:1703.04691*, 2017.

[4] Michael Bruen and Jianqing Yang. Functional networks in real-time flood forecasting—a novel application. *Advances in Water Resources*, 28(9):899–909, 2005.

[5] Jonas Gehring, Michael Auli, David Grangier, Denis Yarats, and Yann N Dauphin. Convolutional sequence to sequence learning. *arXiv preprint arXiv:1705.03122*, 2017.

[6] Nipuni Hettiarachchi and Rohana Thilakumara. Water level forecasting and flood warning system a neuro-fuzzy approach. *International Journal of Applied Engineering Research*, 9, 01 2015.

[7] CE Imrie, S Durucan, and A Korre. River flow prediction using artificial neural networks: generalisation beyond the calibration range. *Journal of hydrology*, 233(1-4):138–153, 2000.

[8] Ibrahem Kandel and Mauro Castelli. The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express*, 2020.

[9] Yoon Kim. Convolutional neural networks for sentence classification. *arXiv preprint arXiv:1408.5882*, 2014.

[10] ÖZGÜR KİŞİ. Daily river flow forecasting using artificial neural networks and auto-regressive models. *Turkish Journal of Engineering and Environmental Sciences*, 29(1):9–20, 2005.

[11] Xuan-Hien Le, Hung Viet Ho, Giha Lee, and Sungho Jung. Application of long short-term memory (lstm) neural network for flood forecasting. *Water*, 11(7):1387, 2019.

[12] David R Legates and Gregory J McCabe Jr. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water resources research*, 35(1):233–241, 1999.

[13] Amir Mosavi, Pinar Ozturk, and Kwok-wing Chau. Flood prediction using machine learning models: Literature review. *Water*, 10(11):1536, 2018.

[14] Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*, 2016.

[15] Rabindra K Panda, Niranjan Pramanik, and Biplab Bala. Simulation of river stage using artificial neural network and mike 11 hydrodynamic model. *Computers & Geosciences*, 36(6):735–745, 2010.

[16] Dario Rethage, Jordi Pons, and Xavier Serra. A wavenet for speech denoising. pages 5069–5073, 2018.

[17] A.R. Rao R.S. Govindaraju. Artificial neural networks in hydrology. page 04014006, 2000.

[18] Astrid Tijskens, Hans Janssen, and Staf Roels. Optimising convolutional neural networks to predict the hygrothermal performance of building components. *Energies*, 12(20):3966, 2019.

[19] Tawatchai Tingsanchali and Mahesh Raj Gautam. Application of tank, nam, arma and neural network models to flood forecasting. *Hydrological Processes*, 14(14):2473–2487, 2000.

[20] Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, and Chengqi Zhang. Graph wavenet for deep spatial-temporal graph modeling. *arXiv preprint arXiv:1906.00121*, 2019.