Green Roof Hydrological Modelling With GRU And LSTM Networks

Haowen Xie (✉ Xhwhhu@outlook.com)
Hohai University  https://orcid.org/0000-0002-3632-8070

Randall Mark
University of Copenhagen

Kwok-wing Chau
Hong Kong Polytechnic University

Research Article

Keywords: Green Roof, Machine Learning, LSTM, GRU, Information Entropy

Posted Date: November 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-922451/v1

License: ☛ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published at Water Resources Management on February 3rd, 2022. See the published version at https://doi.org/10.1007/s11269-022-03076-6.
Green roof hydrological modelling with GRU and LSTM networks

Haowen Xie\textsuperscript{1}\textsuperscript{*}; Mark Randall\textsuperscript{2}\textsuperscript{*}; Kwok-wing Chau\textsuperscript{3}

\textsuperscript{1}Hohai University, Nanjing, 210098, China

\textsuperscript{2}Department of Geosciences and Natural Resources Management, University of Copenhagen, Frederiksberg, Denmark

\textsuperscript{3}Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hong Kong, China

Abstract

Green Roofs (GRs) are becoming more popular as a low-impact building option. They have the potential to minimize peak stormwater runoff while also increasing the quality of runoff from buildings. Improvement of hydrological models for the simulation of GRs will aid design of individual roofs as well as city scale planning that relies on the predicted impacts of widespread GR implementation.

Machine learning (ML) has exploded in popularity in recent years, owing to considerable increases in processing power and data availability. However, there are no studies focusing on the use of ML in hydrological simulation of GRs.

We focus on two types of ML-based model: long short-term memory (LSTM) and gated recurrent unit (GRU) in modelling hydrological performance of GRs, with sequence input and a single output (SISO), and synced sequence input and output
(SSIO) architectures.

According to the results of this paper, LSTM and GRU are useful tools for the modelling of GRs. As the time window length (memory length, time step length of input data) increases, SISO appears to have a higher overall forecast accuracy. SSIO delivers the best overall performance, when the SSIO is close to, or even exceeds, the maximum window size.

**Keywords:** Green Roof, Machine Learning, LSTM, GRU, Information Entropy

**Highlights**

- For the information entropy, the methods contain more information with the increase of window size, and SSIO has the most information over all events. GRU on the whole will contain more information than LSTM.

- SSIO has the best overall performance which does not need window size, and SISO's overall forecast accuracy improves as the duration of the temporal frame grows.

- LSTM and GRU are both useful in modelling GRs, but GRU is better than LSTM.

**1. Introduction**

Traditional stormwater infrastructure promotes rapid runoff from impervious surfaces to receiving water bodies. Urban water infrastructure has been studied extensively with the growth of sustainable development, and stormwater has
increasingly become regarded as a resource. However, many sustainable solutions have a huge footprint and are not viable in urban settings. Green roofs (GRs) have been used to alleviate stormwater management issues in numerous places around the world (Czemiel Berndtsson, 2010, Shafique and Kim, 2017). GRs provide a number of ecosystem services, e.g. biodiversity support, extension of hard-roof life time, reduction in stormwater runoff and noise, improved building insulation, lowering of air temperatures, however the final effect of implementation depends on largely on local context and GR design (Berardi et al. , 2014, Getter and Rowe, 2006, Kolokotsa et al. , 2013, Lepp, 2008, Oberndorfer et al. , 2007).

In this regard, GRs hydrological simulation employs a variety of models. Curve methods (CM), linear/non-linear storage reservoirs (LSR), single reservoir models (SR), and physical models (PM) are four methods accessible for modelling a green roof in general (Daniel Roehr, 2010, Emmanuel Berthier, 2011, Hilten et al. , 2008, Li and Babcock, 2015, Rasmussen, 2006, Soulis et al. , 2017, Versini et al. , 2015, Xie and Liu, 2020, Xie et al. , 2020).

CM is a purely empirical method that relies on the statistical analysis of measured data (Daniel Roehr, 2010, Rasmussen, 2006). Green roofs are treated as a mixture of storage reservoirs in LSR, which typically represent different GRs layers (Emmanuel Berthier, 2011, Kasmin et al. , 2010). SR are based on the water balance equation in a reservoir (Stovin et al. , 2013, Yang et al. , 2015). Finally, PM are concerned with simulating real-world situations and processes (Palla et al. , 2009, She and Pang, 2010, Sun et al. , 2013). In Xie, Wu (2020) and Xie and Liu (2020), the predictions of a non-
linear storage reservoirs model, Storm Water Management Model (SWMM), and a physical process model (HYDRUS-1D) were discussed. Both models were compared against measured data obtained from a series of laboratory experiments. It was concluded that after sufficient debugging, the two models can well simulate the real situation on the index of the total runoff and other event scales, but the simulation of runoff process is not satisfactory, and cannot fit the actual runoff process well. To solve this problem, we use Machine learning (ML) in this study.

ML has been increasingly popular in different disciplines in recent years, owing to considerable increases in computer power and data availability. It has been demonstrated in various studies that a machine learning algorithm could learn from data (LeCun et al., 2015, Mitchell et al., 2003). ML models for streamflow prediction have been widely utilized in hydrology (Dawson and Wilby, 1998, Hettiarachchi et al., 2005, Srinivasulu and Jain, 2006, Sudheer et al., 2002, Young et al., 2017). Recurrent neural network (RNN) methods have recently been utilized for rainfall-runoff (RR) modelling (Taver et al., 2015).

With its built in loops, RNN is capable of collecting long-term data dependencies (Informatik et al., 2003, LeCun, Bengio, 2015, Zhang et al., 2018). However, it may also experience disappearing and exploding gradient problems when it comes to network training (Informatik, Bengio, 2003).

Gradient vanishing indicates that parameters in the hidden layer near the output layer are updated normally, whereas parameters in the hidden layer near the input layer are updated slowly. This is equivalent to the first few hidden layers acting as a
mapping layer for input values, and only the next few hidden layers are really learning. As a result, the learning speed of the neural network is slowed down and learning ability is impaired. In the general case of gradient explosion, the initial weight value is given a relatively large value, resulting in a particularly large change in the updating of hidden layer parameters near the input layer, while the updating of hidden layer parameters near the output layer is relatively normal. Whether gradient vanishing or gradient explosion occurs, this is not the desired result in the training of a neural network, because it will lead to the instability of the gradient change, resulting in poor training and learning results (Hochreiter, 1998). Long short-term memory (LSTM) networks have been created to achieve this goal by incorporating cell state and gating mechanisms into standard RNNs (Gers et al., 2000, Hochreiter, 1998). LSTM network's gates handle the choice of whether to forget or remember information by storing mistakes in memory, which prevents error signal decay (Hochreiter and Schmidhuber, 1997). In other words, LSTM's gates aid in the long-term preservation of states and short-term dependence. The training procedure for building LSTM networks, however, takes a considerable period of time due to its complicated structure. GRU networks were introduced as a simpler version of LSTM networks designed to shorten the training process (Cho et al., 2014).

The advantage of RNNs over standard ANNs is due in part to its sequential regime of operation, which is different from fixed-size networks. Different designs, such as SISO and SSIO, can be employed depending on the network's usage. Unlike the SSIO design, which depends on the LSTM structure to capture lengthy relationships, the
SISO architecture necessitates setting a constant window size (Li et al., 2020).

When a fixed window size is chosen, LSTM and GRU are forced to confine dependencies to the size of the chosen window. Due to the transference of hidden states from earlier time steps, the SSIO architecture, on the other hand, is capable of capturing long-term dependencies on its own. In other words, if the user wants to utilize a fixed window regime, there is no need to employ LSTM and GRU for hydrology modelling because the LSTM network is not required to provide the model with a fixed window size. Furthermore, determining the window size necessitates a thorough understanding of the watershed's reaction to rainfall events (Hu et al., 2018, Kratzert et al., 2018, Kratzert et al., 2019, Yuan et al., 2018).

In GR modeling, Asadi et al. (2020) utilized machine learning to establish a link between LST and different urban characteristic factors at the same time, and used an artificial neural network to model GRs and their possible mitigating effects on an urban heat island with a case study in Austin, Texas. Erdemir and Ayata (2017) proposed a model for predicting temperature drop on a green roof using an artificial neural network (ANN). Tsang and Jim (2016) used meteorological data to model soil moisture variations and design an effective irrigation plan. In that study, the researchers used artificial intelligence algorithms made up of artificial neural networks and fuzzy logic. However, to the authors’ knowledge, there are no articles focusing on using ML in GRs’ hydrological simulation.

In this paper we make a comparison of LSTM and GRU models with different architectures to verify which methods and architectures are capable of improving the
accuracy of the hydrological modelling of GRs. With the use of the developed model, we aim to answer the following three questions:

i. Which method has better performance in GR hydrological modelling, LSTM or GRU;

ii. Which architecture (SSIO or SISO) of both LSTM and GRU has better performance in GR hydrological modelling;

iii. Which window size (SSIO) of both LSTM and GRU demonstrates better performance in GR hydrological modelling.

2. Methods and data

2.1 Data pre-processing

The Blue Green Wave of Champs-sur-Marne (France) is the Greater Paris Area's largest green roof (1 hectare). In that study, Versini (2019) gathered flow discharge and precipitation data with a resolution of 30 seconds. Data from modified measuring sensors was gathered for 78 days between February and May 2018 in that investigation. Versini et al. (2019) published the open data, which is used in this work. There are no missing values in the rainfall dataset: 7 March (9mm), 11 March (9.7mm), 17 March (7.5mm), 27 and 28 March (13.9mm), 9 April (9.6mm), and 29 and 30 April (9.6mm) (23.5mm).

A total of 22,442 rainfall and runoff records from 6 rainfall-runoff occurrences were chosen.

Model training is based on 9 March (9 mm), 11 March (9.7 mm), 17 March (7.5
mm), and 27 and 28 March (13.9 mm) events (with 16,714 rainfall and runoff records), whereas model testing is based on 9 April (9.6 mm), 29 and 30 April (23.5 mm) events (with 5,728 rainfall and runoff records).

Eq. (2). is used to standardize rainfall and runoff data. The normalized data are in the [0, 1] range.

\[
X_{\text{norm}} = \frac{X_t - X_{\min}}{X_{\max} - X_{\min}} \quad (1)
\]

where \( X_{\text{norm}} \), \( X_t \), \( X_{\min} \) and \( X_{\max} \) are the normalized, observed, minimum and maximum values, respectively, of target variable (rainfall, soil moisture or runoff).

2.2 Problem formulation

We utilize \( Q_t \) to indicate the flow rate at the outflow of a GR at time step \( t \), which is the quantity of interest of this issue, to define a GR hydrological modelling job in a data-driven environment. Similarly, rainfall and runoff records at time step \( t \) in the GR are denoted by \( x_t = \{x_{it}, i = 1,2, \ldots \} \) and \( y_t = \{y_{it}, i = 1,2, \ldots \} \), respectively.

In addition to precipitation, runoff is influenced by other factors such as topography, seasons, and other variables. However, as this research involves simulation with a high temporal resolution (every 30 sec), rainfall and soil moisture, are assumed to be the dominant drivers of runoff. To summarize, runoff modelling is concerned with determining the regression connection between output runoff and input rainfall.
2.3 Principle of Neural Network

2.3.1 LSTM

Hochreiter and Schmidhuber (1997) suggested LSTM to cope with exploding and vanishing gradient issues. Figure 2 depicts the structure of the LSTM. A cell state $c_t$, an input gate $i_t$, a forget gate $f_t$, a cell gate $g_t$, and an output gate $o_t$ make up the LSTM unit.

The updated hidden state $h_t$ is computed by the following equations for each time step $t$ given the input vector $X_t$ (including $x_t$, $y_t$), prior hidden cell state $h_{t-1}$, and previous cell state $c_{t-1}$:

\[
\begin{align*}
        i_t &= \sigma(W_{ii}X_t + b_{ii} + W_{hi}h_{t-1} + b_{hi}) \\
        f_t &= \sigma(W_{if}X_t + b_{if} + W_{hf}h_{t-1} + b_{hf}) \\
        g_t &= \tanh(W_{ig}X_t + b_{ig} + W_{hg}h_{t-1} + b_{hg}) \\
        o_t &= \tanh(W_{io}X_t + b_{io} + W_{ho}h_{t-1} + b_{ho}) \\
        c_t &= f_t \cdot c_{t-1} + i_t \cdot g_t \\
        h_t &= o_t \cdot \tanh(c_t)
\end{align*}
\]

where $\sigma(\cdot)$ is a sigmoid function, and $\cdot$ stands for Hadamard products. Weight matrices are all $W$’s, while bias matrices are all $b$’s. Future time steps should not influence past time steps because real-world time series data are utilized. As a result, the bidirectional mechanism is not used in this research. A single directional LSTM network is used in both designs.
2.3.2 GRU

Figure 3 depicts the GRU cell’s unique structure. In the GRU, the hidden state \( h_t \) and cell state \( c_t \) blend into one. The updated gate \( z_t \) and the reset gate \( r_t \) are two control gates in the GRU cell.

The updated gate \( z_t \) is used to regulate the amount of state information \( h_{t-1}(c_{t-1}) \) delivered into the current time step \( t \) from the previous time step \( t-1 \). The greater the value of the updated gate, the more prior time step state information is brought in.

The reset gate \( r_t \) determines how much data from the previous state is put into the current candidate set \( \tilde{c}_t \). The less information from the previous state is written, the smaller the reset gate \( r_t \) is.

Updated equations in the construction of the GRU cell are determined as follows and also illustrated in Fig. 3. The parameters are the same as those in Section 2.3.1.

\[
r_t = \sigma(W_r x_t + b_l + W_h h_{t-1} + b_h)
\]

(8)
\[ z_t = \sigma(W_{iz} x_t + b_{iz} + W_{hz} h_{t-1} + b_{hz}) \]  
\[ \tilde{c}_t = \tanh(W_{ic} x_t + W_{hc}(r_t * h_{t-1}) + b_c) \]  
\[ c_t = (1 - z_t) * c_{t-1} + z_t * \tilde{c}_t \]  
\[ h_t = c_t \]  

Fig. 3. The structure of a GRU cell.

### 2.3.3 LSTM and GRU Architectures
This paper focuses on two architectures, namely SISO and SSIO (see Fig. 4). Input and output of SSIO have the same length and do not need to be fixed. But the input window size becomes an additional hyperparameter that must be tuned for SISO. As a result, the architecture is addressed briefly in the introduction without going into detail in this study. Since no previous study had focus on methods in GRs modelling and the interest in this paper is only in estimating runoff based on precipitation, two architectures, namely SISO and SSIO, are compared (Hu, Wu, 2018, Kratzert, Klotz, 2018, Kratzert, Klotz, 2019, Yuan, Chen, 2018).

2.4 Model development

2.4.1 SSIO model of LSTM and GRU

Because LSTM and GRU networks are very similar, identical hyperparameters are
developed for them.

A trial and error method is used to tune the parameters in this study (Liang et al., 2018). It is found that, after a significant number of tests, only one layer is superior to using a multi-layer network to mimic RR interactions in the investigated area. At each time step, the suggested networks have a 5 neuron input layer, a 20 neuron hidden layer, and a single neuron output layer.

The batch size and epoch size are both set at 64 and 500. The process of sending all data into the network to complete an iterative computation is referred to as an epoch. All data in our model are iterated 500 times. The ‘Relu' activation function is selected. RMSprop is the optimizer utilized in this study.

2.4.2 SISO model of LSTM and GRU

For the SISO model, four alternative window sizes (memory lengths), namely 20 steps, 120 steps, 240 steps, and 360 steps, are evaluated. These options correspond to 5, 30, 60, and 90 minutes, respectively. The progressive options are intended to demonstrate the GR hydrological modeling's long-term dependence cascade. The SISO models' additional hyperparameters are the same as those of the SSIO models in section 2.4.1. Hyperparameters in models with varied time steps and lead times are the same.

2.5 Evaluation metrics

The total runoff ratio ($R_V$), Nash–Sutcliffe efficiency (NSE), root mean square error (RMSE), and mean absolute error (MAE) are employed as error indicators.
\[ R_v = \frac{\sum_{t=1}^{t=t_F} Q_{\text{sim}}(t) \Delta t}{\sum_{t=1}^{t=t_F} Q_{\text{mea}}(t) \Delta t} \]  

(13)

\[ \text{NSE} = 1 - \frac{\sum_{t=1}^{t=t_F} (Q_{\text{mea}}(t) - Q_{\text{sim}}(t))^2}{\sum_{t=1}^{t=t_F} (Q_{\text{mea}}(t) - Q_{\text{Amea}})^2} \]  

(14)

\[ R_{QM} = \frac{Q_{M\text{sim}}}{Q_{M\text{mea}}} \]  

(15)

\[ RMSE = \left( \frac{\sum_{t=1}^{t=t_F} [Q_{\text{mea}}(t) - Q_{\text{sim}}(t)]^2}{\sum_{t=1}^{t=t_F} 1} \right)^{1/2} \]  

(16)

\[ MAE = \frac{\sum_{t=1}^{t=t_F} |Q_{\text{sim}}(t) - Q_{\text{mea}}(t)|}{\sum_{t=1}^{t=t_F} 1} \]  

(17)

The complexity of hydrological processes and a lack of understanding of hydrological prototypes are usually responsible for the diversity of models (Xia et al., 1997, Xie and Liu, 2020). The model's information entropy measure \( I(Q_i) \) can measure the amount of information it contains.

\[ I(Q_i) = -K \sum_{j=1}^{S} y_i \ln y_i \quad (j \in J; K = \frac{1}{\ln 2}) \]  

(18)

\[ y_i(k) = \frac{1}{1 + \Delta_i Q(k)} \frac{1}{\sum_{i=1}^{n} \left( \frac{1}{1 + \Delta_i Q(k)} \right)} \]  

(19)

\[ \Delta_i Q(k) = |Q_i(k) - Q(k)| \]  

(20)

where \( y_i \) is the normalised residual order; \( \Delta_i Q(k) \) is model i’s residuals; \( \{Q(k)\}, k = 1, 2, \cdots, n, \) is GRs' observed runoff sequence; \( \{Q_i(k)\}, k = 1, 2, \cdots, n, \) is the estimated runoff flow rate sequence for hydrological model \( i. \)
3. Results

Table 2 displays the simulation results. The SSIO model has the best overall performance amongst all techniques examined in this study.

For SISO models, the total prediction accuracy improves as the duration of the time frame grows. As the time frame grows from 20 to 360 steps, NSE, MAE, and RMSE show a clear progressive increase. This behavior implies that runoff at the investigated GR is truly long-term dependent on prior rainfall history; hence, to simulate on a fine temporal scale, architectures that can retain long-term memory are necessary.

From Table 2, $R_V$ is less than 1, which means that none of the methods overestimate the events, and the values of SSIO and SISO 360 are closest to the measured values. The results of SISO 20-240 get better with the increase of window size.

Table 2 shows the information entropy metric $I(Q)$. The greater the value of $I(Q)$, the more information the model includes. Nevertheless, the value will change depending on the event. The $I(Q)$ of the event on 29 and 30 April (23.5mm) are more than that of the event on 9 April (9.6mm) for all methods, which means events with higher rainfall tend to have more information. For LSTM and GRU, $I(Q)$ of SISO 20-360 gets better with the increase of window size and results of SSIO are close to the best values of SISO, demonstrating that methods contain more information with the increase of window size, and SSIO has the most information overall events. Comparing LSTM and GRU models, they are close in the event on 29 and 30 April (23.5mm), and GRU is better than LSTM in the event on 9 April (9.6mm). It can be
concluded that GRU on the whole contains more information than LSTM.

**Table 2.** Summary of indexes (a) Results of LSTM; (b) Results of GRU.

(a)

| Test events | Test | NSE  | Rv   | RMSE  | MAE  | I (Q) |
|-------------|------|------|------|-------|------|-------|
| 9 April     |      |      |      |       |      |       |
| (9.6mm)     |      |      |      |       |      |       |
| 9 April     | SSIO | 0.939| 0.928| 0.002 | 0.001| 5.337 |
|             | SISO 20 | 0.885| 0.859| 0.006 | 0.005| 5.305 |
|             | SISO 120 | 0.947| 0.890| 0.003 | 0.002| 5.314 |
|             | SISO 240 | 0.938| 0.929| 0.002 | 0.001| 5.371 |
|             | SISO 360 | 0.944| 0.977| 0.001 | 0.001| 5.343 |
| 29 and 30   | SSIO | 0.980| 0.972| 0.001 | 0.001| 5.914 |
| April       | SISO 20 | 0.847| 0.626| 0.006 | 0.004| 5.832 |
| (23.5mm)    | SISO 120 | 0.926| 0.833| 0.002 | 0.001| 5.889 |
|             | SISO 240 | 0.950| 0.932| 0.001 | 0.001| 5.918 |
|             | SISO 360 | 0.992| 0.965| 0.001 | 0.000| 5.963 |

(b)

| Test events | Test | NSE  | Rv   | RMSE  | MAE  | I (Q) |
|-------------|------|------|------|-------|------|-------|
| 9 April     |      |      |      |       |      |       |
| (9.6mm)     |      |      |      |       |      |       |
| 9 April     | SSIO | 0.953| 0.941| 0.001 | 0.000| 5.757 |
|             | SISO 20 | 0.873| 0.811| 0.002 | 0.001| 5.633 |
|             | SISO 120 | 0.965| 0.914| 0.001 | 0.001| 5.718 |
|             | SISO 240 | 0.972| 0.911| 0.001 | 0.000| 5.674 |
|             | SISO 360 | 0.997| 0.950| 0.001 | 0.000| 5.739 |
|        | SSIO  | April  |
|--------|-------|--------|
|        |       | (23.5mm) |
| 29 and 30 | 0.995 | 0.971  | 0.001 | 0.000 | 5.966 |
|        | SISO 20 | 0.904  | 0.913 | 0.001 | 0.001 | 5.863 |
| SISO 120 | 0.914  | 0.929  | 0.001 | 0.001 | 5.990 |
| SISO 240 | 0.949  | 0.968  | 0.001 | 0.001 | 5.997 |
| SISO 360 | 0.997  | 0.951  | 0.001 | 0.000 | 5.996 |
Fig. 5 Predictions from LSTM models vs observed values
Fig. 6 Predictions from GRU models vs observed values
Figures 5 and 6 show LSTM and GRU model output for the event on 29 and 30 April (23.5mm) compared to the observed values. It can be seen that all approaches are successful in identifying a similar runoff pattern. However, as the length of input memory increases, increasingly better performance is observed. SISO's simulation performance approaches that of SSIO model when the time window size approaches 360 steps. From Figures 5 and 6, the peaks of all modelling runoff values are smaller than those of the observed values, which also indirectly confirms the result that $R_V$ of all methods are not greater than 1. Moreover, the peak values of all modelling runoff values are earlier than those of the measured values, and the values of SSIO and SISO 360 are the closest to the measured values. Results of SISO 20-240 get better with the increase of window size.

4. Conclusion

This paper is the first research, to the authors’ knowledge, focusing on the use of ML in GR hydrological simulation. The performances of LSTM and GRU models with SISO, SSIO architectures in modelling the hydrological performance of GRs are examined in this study.

According to results of this paper, LSTM and GRU are useful tools in modelling GRs. The total forecast accuracy of SISO improves as the length of the temporal window rises. SSIO has the best overall performance and the performance of SSIO is close to, or even exceeds, the maximum window size. For the volume of the total
runoff, LSTM and GRU do not overestimate runoff and ML methods could be integrated into the existing hydrologic models. For the information entropy, the methods contain more information with the increase of window size, and SSIO has the most information over all events. GRU on the whole will contain more information than LSTM.

In conclusion, LSTM and GRU are both useful in modelling GRs, but GRU is better than LSTM. For architectures, SSIO has the best overall performance which does not need window size, and SISO's overall forecast accuracy improves as the duration of the temporal frame grows.

This study has demonstrated the potential of ML methods to improve the simulation of GRs, which may help pave the way for related tools to be implemented in existing hydrologic models. Better representation of individual stormwater management elements including GRs in hydrologic models will aid the design of individual GRs and improve the accuracy of large scale models that support city planning.

5. References

Asadi A, Arefi H, Fathipoor H. Simulation of green roofs and their potential mitigating effects on the urban heat island using an artificial neural network: A case study in Austin, Texas. Advances In Space Research. 2020;66:1846-62.
Berardi U, GhaffarianHoseini A, GhaffarianHoseini A. State-of-the-art analysis of the environmental benefits of green roofs. Appl Energ. 2014;115:411-28.
Cho K, Van Merrienboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, et al. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. Computer ence. 2014.
Czemiel Berndtsson J. Green roof performance towards management of runoff water quantity and quality: A review. Ecological Engineering. 2010;36:351-60.
Daniel Roehr YK. Runoff Reduction Effects of Green Roofs in Vancouver, BC, Kelowna, BC, and Shanghai. Canadian Water Resources Journal 2010;35:15.
Dawson CW, Wilby R. An artificial neural network approach to rainfall-runoff modelling. Hydrolog Sci J. 1998;43:47-66.

Emmanuel Berthier DR, Bernard de Gouvello. Simulation of green roof hydrological behavior with a reservoir model. 12nd International Conference on Urban Drainage. Porto Alegre/Brazil2011.

Erdemir D, Ayata T. Prediction of temperature decreasing on a green roof by using artificial neural network. Appl Therm Eng. 2017;112:1317-25.

Gers FA, Schmidhuber J, Cummins F. Learning to forget: Continual prediction with LSTM. Neural Computation. 2000;12:2451-71.

Getter KL, Rowe DB. The role of extensive green roofs in sustainable development. Hortscience. 2006;41:1276-85.

Hettiarachchi P, Hall MJ, Minns AW. The extrapolation of artificial neural networks for the modelling of rainfall-runoff relationships. J Hydroinformatics. 2005;7:291-6.

Hilten RN, Lawrence TM, Tollner EW. Modeling stormwater runoff from green roofs with HYDRUS-1D. Journal of Hydrology. 2008;358:288-93.

Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. International Journal Of Uncertainty Fuzziness And Knowledge-Based Systems. 1998;6:107-16.

Hochreiter S, Schmidhuber J. Long short-term memory. Neural Computation. 1997;9:1735-80.

Hu C, Wu Q, Li H, Jian S, Li N, Lou Z. Deep Learning with a Long Short-Term Memory Networks Approach for Rainfall-Runoff Simulation. Water. 2018;10.

Informatik F, Bengio Y, Frasconi P, Schmidhuber J. Gradient Flow in Recurrent Nets: the Difficulty of Learning Long-Term Dependencies. A Field Guide to Dynamical Recurrent Neural Networks. 2003.

Kasmin H, Stovin VR, Hathway EA. Towards a generic rainfall-runoff model for green roofs. Water science and technology : a journal of the International Association on Water Pollution Research. 2010;62:898-905.

Kolokotsa D, Santamouris M, Zerefos SC. Green and cool roofs’ urban heat island mitigation potential in European climates for office buildings under free floating conditions. Solar Energy. 2013;95:118-30.

Kratzert F, Klotz D, Brenner C, Schulz K, Herrnegger M. Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. Hydrol Earth Syst Sc. 2018;22:6005-22.

Kratzert F, Klotz D, Shalev G, Klambauer G, Hochreiter S, Nearing G. Benchmarking a Catchment-Aware Long Short-Term Memory Network (LSTM) for Large-Scale Hydrological Modeling2019.

LeCun Y, Bengio Y, Hinton G. Deep Learning. Nature. 2015;521:436-44.

Lepp NW. Planting Green Roofs and Living Walls. J Environ Qual. 2008;37:2408-.

Li W, Kiaghadi A, Dawson C. Exploring the best sequence LSTM modeling architecture for flood prediction. Neural Computing & Applications. 2020.

Li Y, Babcock RW, Jr. Modeling Hydrologic Performance of a Green Roof System with HYDRUS-2D. J Environ Eng. 2015;141.

Liang C, Li H, Lei M, Du, Qingyun. Dongtng Lake Water Level Forecast and Its Relationship with the Three Gorges Dam Based on a Long Short-Term Memory Network. Water. 2018;10:1389.

Mitchell T, Buchanan B, DeLong G, Dietterich T, Rosenbloom P, Waibel A. Machine Learning. Annual Review of Computer Science. 2003;4:417-33.

Najafzadeh M. Neurofuzzy-Based GMDH-PSO to Predict Maximum Scour Depth at Equilibrium at Culvert Outlets. Journal of Pipeline Systems Engineering and Practice. 2015;7:06015001.

Oberndorfer E, Lundholm J, Bass B, Coffman RR, Doshi H, Dunnett N, et al. Green roofs as urban
ecosystems: Ecological structures, functions, and services. Bioscience. 2007;57:823-33.
Palla A, Gnecco I, Lanza LG. Unsaturated 2D modelling of subsurface water flow in the coarse-grained porous matrix of a green roof. Journal Of Hydrology. 2009;379:193-204.
Rasmussen Ca. HYDROLOGIC BEHAVIOR OF VEGETATED ROOFS. JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION. 2006;42:1261–74.
Shafique M, Kim R. Retrofitting the Low Impact Development Practices into Developed Urban areas Including Barriers and Potential Solution. Open Geosciences. 2017;9:240-54.
She N, Pang J. Physically Based Green Roof Model. Journal Of Hydrologic Engineering. 2010;15:458-64.
Soulis KX, Valiantzas JD, Ntoulas N, Kargas G, Nektarios PA. Simulation of green roof runoff under different substrate depths and vegetation covers by coupling a simple conceptual and a physically based hydrological model. Journal of environmental management. 2017;200:434-45.
Srinivasulu S, Jain A. A comparative analysis of training methods for artificial neural network rainfall–runoff models. Applied Soft Computing. 2006;6:295-306.
Stovin V, Poë S, Berretta C. A modelling study of long term green roof retention performance. Journal of environmental management. 2013;131:206-15.
Sudheer KP, Gosain AK, Ramasastri KS. A data-driven algorithm for constructing artificial neural network rainfall-runoff models. Hydrol Process. 2002;16:1325-30.
Sun T, Bou-Zeid E, Wang Z-H, Zerba E, Ni G-H. Hydrometeorological determinants of green roof performance via a vertically-resolved model for heat and water transport. Building And Environment. 2013;60:211-24.
Taver V, Johannet A, Borrell-Estupina V, PISTRE S. Feed-forward vs recurrent neural network models for non-stationarity modelling using data assimilation and adaptivity. Hydrolog Sci J. 2015;60:1242-65.
Tsang SW, Jim CY. Applying artificial intelligence modeling to optimize green roof irrigation. Energy And Buildings. 2016;127:360-9.
Versini P-A. Blue Green Wave hydrological data [Data set]. Zenodo. 2019.
Versini PA, Ramier D, Berthier E, de Gouvello B. Assessment of the hydrological impacts of green roof: From building scale to basin scale. Journal of Hydrology. 2015;524:562-75.
Xia J, Huang GH, Bass B. Combination of Differentiated Prediction Approach and Interval Analysis for the Prediction of Weather Variables Under Uncertainty. Journal of environmental management. 1997;49:95-106.
Xie H, Liu J. A modeling study of the interflow in the green roof. Urban for Urban Gree. 2020;54.
Xie H, Wu Y, Wang L, Luo W, Zhou W, Zhou H, et al. Comparing simulations of green roof hydrological processes by SWMM and HYDRUS-1D. Water Supply. 2020;20:130-9.
Yang W-Y, Li D, Sun T, Ni G-H. Saturation-excess and infiltration-excess runoff on green roofs. Ecological Engineering. 2015;74:327-36.
Young C-C, Liu W-C, Wu M-C. A physically based and machine learning hybrid approach for accurate rainfall-runoff modeling during extreme typhoon events. Applied Soft Computing. 2017;53:205-16.
Yuan X, Chen C, Lei X, Yuan Y, Adnan RM. Monthly runoff forecasting based on LSTM-ALO model. Stoch Env Res Risk A. 2018;32:2199-212.
Zhang J, Zhu Y, Zhang X, Ye M, Yang J. Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas. Journal Of Hydrology. 2018;561:918-29.