

Demonstration of the impacts of anti-sedimentation techniques on Japanese reservoir siltation via mass data ANN analysis

Tobias Landwehr^{a,*}, Sameh Ahmed Kantoush^{id}^b, Daisuke Nohara^c, Tetsuya Sumi^b and Claudia Pahl-Wostl^a

^aInstitute for Environmental Systems Research, Osnabrück University, Barbarastraße 12, 49076 Osnabrück, Germany

^bWater Resources Research Center, Disaster Prevention Research Institute, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

^cKajima Technical Research Institute Tobitakyu 2-19-1, Chofu 182-0036, Japan

*Corresponding author. E-mail: tobias.landwehr@uos.de

 SAK, 0000-0003-0919-5097

ABSTRACT

Reservoirs have been installed as long-term assets to guarantee water and energy security for decades, if not centuries. However, the effect of siltation undermines reservoirs' sustainability because it significantly reduces the reservoirs' original capacity. Extreme events such as typhoons, floods and droughts are posited to have extreme impacts on sediment inflow and deposition in reservoirs. The same holds true for ISMTs (implemented sediment management technologies), such as dredging, spilling and bypassing. However, the large-scale analysis of their effects on reservoir sedimentation progression, recovery and development was not feasible due to data scarcity and technological restrictions. The present paper closes this information gap by conducting a GRU (gated recurrent unit) neural network analysis of 1,224 Japanese reservoirs, for which the sedimentation, local precipitation, extreme events and ISMTs were monitored between 2000 and 2017. The network reveals the beneficial impacts of dredging, spilling and bypassing. The results also demonstrate the potential of smart management and improved monitoring for sedimentation threat abatement. Thus, foresighted engineering and dedicated governance action in flood and drought scenarios can significantly strengthen the sustainable behavior of key infrastructure elements such as reservoirs.

Key words: Japan, mass data, neural networks, reservoirs, RNN, sediment management

HIGHLIGHTS

- Unique data set: 1,225 dams with 18 years of sediment record each.
- Anti-sediment management notations.
- Artificial neural networks with gated recurrent units as methodology.
- Unique conclusions regarding the efficiency of sediment management methodology.
- Generalized results represent multifaceted types of reservoirs.

1. INTRODUCTION

The number and capacity of hydropower sites and reservoirs are expected to increase within the next decades (Annandale 2013). For example, Zarfl *et al.* (2015) assume that the global quantity of installed dams will double from 2010 to 2030. Although a non-negligible number of proposed reservoirs will not be completed, the overall reservoir trend consists of constant or disproportional growth (Dogmus & Nielsen 2019).

A partial explanation for the construction boom is the increasing siltation in already inaugurated dams, which has reduced the gross reservoir volume per capita (Annandale 2013) and the gross reservoir volume (Oehy 2003; Kantoush & Sumi 2010) (as shown in Figure 1) to the level of the 1970s despite the ever-increasing number of reservoirs. Figure 1 demonstrates clearly that in 2010 global sedimentation was assumed to grow faster than the global gross reservoir volume. This is a giant setback for sustainable long-term renewable energy generation and water supply security.

Studies show that planners, operators and practitioners underestimate the siltation threat, whereas they overestimate their own capacities, namely, that reservoir sedimentation management suffers from optimism bias (Annandale 2013; Flyvbjerg 2016; Schleiss *et al.* 2016; Landwehr *et al.* 2020). Thus, it is no surprise that the implementation and analysis of suitable

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

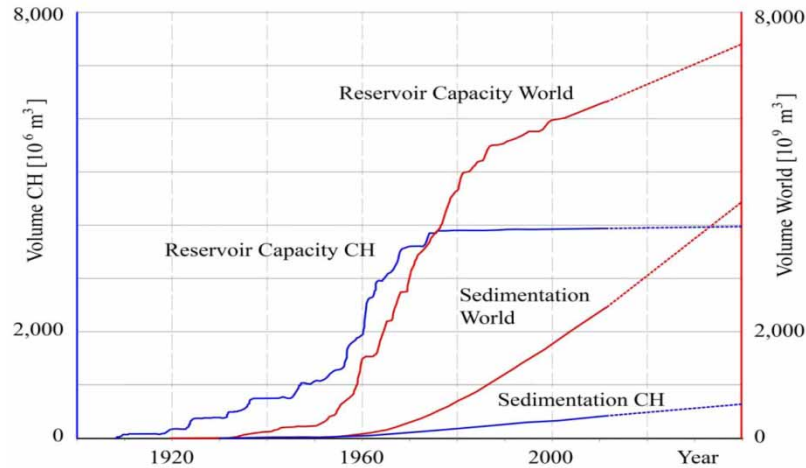


Figure 1 | Assumed global (and Swiss) siltation development according to Oehy 2003; Kantoush & Sumi (2010).

management strategies remain challenging (de Vente *et al.* 2013; Yang 2013; Kantoush & Sumi 2017) despite the availability of manifold elaboration prediction and simulation approaches (Simoes & Yang 2006; Zeleke *et al.* 2013; Omer *et al.* 2015; Ghimire & DeVantier 2016; Hao *et al.* 2017). This scenario prevails regardless of the environment, climate, society and technological level and threatens the designed functionality of reservoirs worldwide (Basson 2009; Schleiss *et al.* 2010; Annandale 2013).

This is alarming, as reservoirs are a junction infrastructure element of the WEF (water–energy–food) nexus, namely, via crop production or hydroelectricity. Hence, reservoirs and their sedimentation are also crucial factors in the realization of highly interlinked sustainable development goals (SDGs) (United Nations 2015; Zhang *et al.* 2016; Pousse & Latouche 2018).

Various technological, management and governance approaches to amend, reduce or reverse sediment accumulation in reservoirs are applied globally to a scant degree compared to the global gross reservoir volume (Morris & Fan 1998; Kondolf *et al.* 2014; Pahl-Wostl 2015; Sumi 2015; Kantoush & Sumi 2017, 2019; Morris 2020). Moreover, the analysis of these approaches has been restricted to local or regional studies (Haregeweyn *et al.* 2013; Pandey *et al.* 2016; Velásquez-Castro *et al.* 2016; Wild *et al.* 2016; Adeogun *et al.* 2018) of single or a few reservoirs due to a lack of siltation data (Schleiss *et al.* 2016). The same holds true for events that are beyond human control, namely, force majeure such as floods, typhoons or earthquakes, which have substantial impacts but the analysis of which is typically local or regional (Lee *et al.* 2006; Vanmaercke *et al.* 2014; Wang *et al.* 2018; Stähly *et al.* 2019).

Nevertheless, those studies do not incorporate the decisive learning factor called experience that surges from large-scale, long-term data. Such data-driven studies and approaches regarding reservoir sedimentation management and event impact have proven to be challenging in recent decades, as few authorities monitor sediment in a regular, overarching manner.

One of the exceptions in data monitoring, however, is found in Japan, which is also one of the leading countries in applied sediment management (Kondolf *et al.* 2014; Auel *et al.* 2016; Kantoush & Sumi 2017). Japanese data collection regarding siltation is comparatively vast in terms of both time and extent (Landwehr *et al.* 2020).

However, siltation is a nonlinear time-series process (Annandale 2013). An emergent tool for nonlinear data analysis and nonlinear data emulation is artificial neural networks (ANNs), the advantages of which include comparatively rapid applicability and vast use case flexibility (Gamboa 2017). Recurrent neural networks (RNNs), which include long short-term memory (LSTM) and gated recurrent unit (GRU) variants, are especially suitable for surveying complicated and intertwined time-series processes (Fu *et al.* 2016; Petneházi 2019; Elsworth & Güttel 2020).

By utilizing RNNs, the present paper seeks to derive evidence from the Japanese dataset that various events or management technologies and actions have retraceable impacts on reservoir siltation. The study aims at deducing these effects from a big long-term data picture to obtain generalized results that might be applied or reproduced globally. In most countries of the world, data are scarce; hence, to pursue global reproducibility, this paper attempts to obtain improved results but not maximum confidence with a highly reduced data input.

The pursuit of general results implies that this analysis does not consider every single dam in a highly specific manner – it is acknowledged that there are huge individual differences. Nevertheless, if the Japanese reservoir variety is regarded as a reflection of the global or supra-regional reservoir variety, the produced results will provide worthwhile insights into global reservoir sustainability.

2. METHODOLOGY

Reservoir siltation data suffer from ambiguity due to the influence of other parameters. Figure 2 presents the absolute sedimentation in comparison to the age of three exemplarily selected reservoirs. It also shows occurrences of (natural) events or implemented sediment management technologies (ISMTs) (from now on, *event categories* are used to refer to both).

Information on how the event categories influence sedimentation is difficult to derive from these few and partially contradictory examples. Thus, additional information is needed; tools that can extract information from a complex web of data, such as ANNs, are mandatory.

The objective of the ANN methodology in the present paper is to train a network on a reduced set of data that also includes notifications on event categories. Based on the data characteristics, the ANN shall finally form a variable that insinuates what the network has learned regarding each event category. Technically, the ANN and data processing is carried out via Tensorflow, Keras and Python in a Jupyter compiler, while making use of standard packages of numpy, pandas or sklearn.

In this study, the network aims at learning the execution of flexible and generalized hindcast emulations on the siltated whole volume for each reservoir. To do so, the network must learn data characteristics for each category of input data. This also holds true for the event categories.

The study shall finally display whether the event categories exert general effects on siltation in reservoirs. It shall determine whether those effects are siltation reducing or enforcing. To do so, a semi-randomly selected variance of original reservoir data is modified and retested on the ANN hindcast emulations with modified data. The modification regards solely the event categories.

The deviation from the hindcast emulations with original event category data creates the desired variable that expresses the effect of each event category. A simplified outline is presented in Figure 3.

The process uses the following *input variables*: *event categories* (variable of which the effect is surveyed), *siltation to whole volume* (variable that is to be affected by the others), *reservoir size* (to roughly distinguish between substantially different reservoir types), *precipitation for each month* (to roughly distinguish between substantially different reservoir types), *year* (as a marker for a continuous time series) and *initial sedimentation to whole volume* (to enable the ANN to create a time-series hindcast). The *output variable* is the: *general deviation of hindcasts that were manipulated by a sole event category from the original hindcast* (to measure the relative effect of an event category on reservoir siltation).

A wide range of methodologies were applied in the stages of the present paper to execute the outline of Figure 3. These included data preparation, network optimization and emulation and analysis of the results. The whole process will be explained subsequently in detail. It is illustrated in Figure 4.

2.1. Base data

The base data correspond to 1,224 reservoirs throughout Japan. The data originate from various files and formats, which were joined in a semiautomated process. Altogether, a compound of data was created, which includes the reservoir size (very small: <0.35 million m³, small: > 0.35 million m³, medium: > 1 million m³, large: > 10 million m³ and very large: > 100 million m³

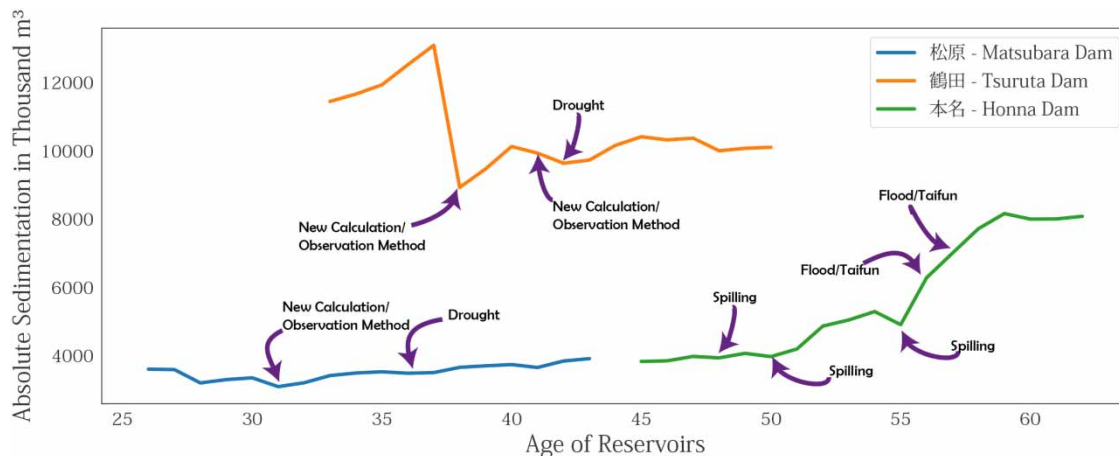


Figure 2 | Three examples show that the impacts of event categories on siltation are not (easily) derivable from a few examples.

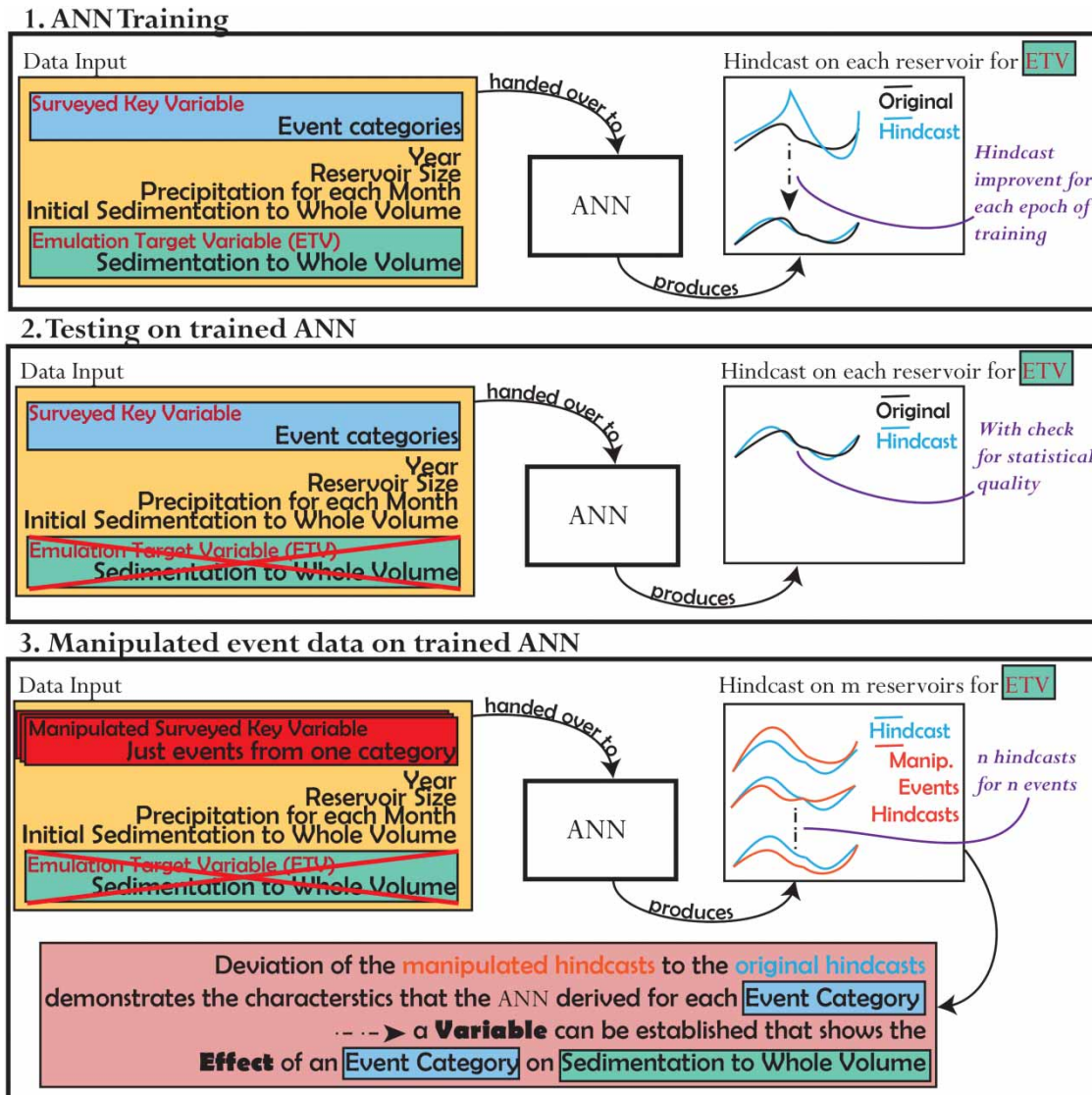


Figure 3 | Three main steps for obtaining the desired information on the impacts of ISMT and events on siltation.

total storage volume [this definition was established for Japan and is not internationally applicable]), siltated whole volume, inauguration year, average monthly precipitation for each month and, most importantly, management and event notifications of the operators. The set also includes information regarding the reservoir purpose, siltated inactive storage, siltated active storage, location and river basin affiliation; additional information was not included in the process of the present paper to further reduce the number of data input variables (see Section 4.5).

The data format is annual and incorporates all years from 2000 to 2017 (*Heisei 12-29* – 平成 1 2 - 2 9 年); hence, each reservoir corresponds to 18 time step time series. Stations of data origin are depicted in [Figure 5](#). The data are structured in three dimensions: reservoirs, time steps and features.

The bulk of the data was provided by the *Kokudo Koutsuu Shou* – 国土交通省 (Japanese Ministry of Land, Infrastructure, Transport and Tourism, MLIT) and the *Doboku Kenkyuu Sho* – 土木研究所 (Public Works Research Institute, PWRI). Only the average monthly precipitation data were obtained from the *Kishouchou* – 気象庁 (Japan Meteorological Agency, JMA).

Hence, the data reflect the general characteristics of Japan's reservoir landscape.

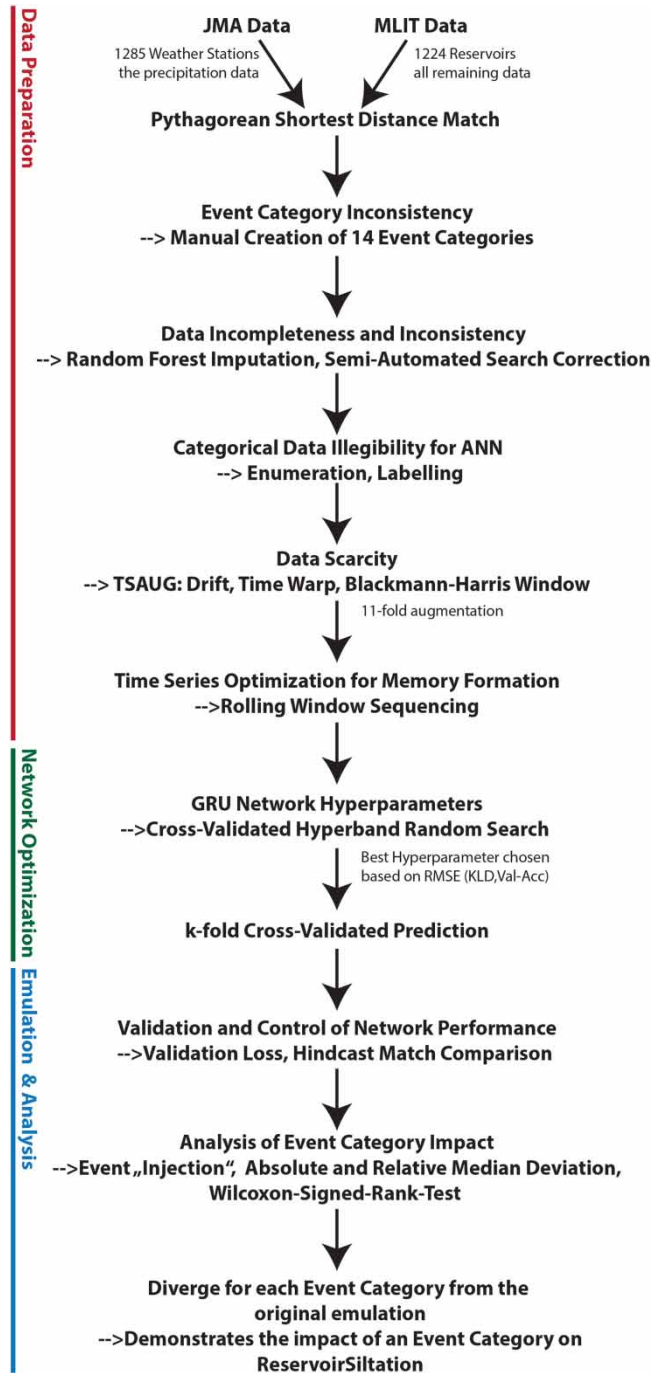


Figure 4 | Representation of the methodological process of the present paper.

2.2. Artificial neural networks

For the obtained data, a nonlinear relation is presumed. ANNs are considered a suitable tool for analysis, pattern detection and data emulation for this type of data behavior (Schmidhuber 2014; Gamboa 2017; Michelucci 2018), especially RNNs (Fu *et al.* 2016; Petneházi 2019; Elsworth & Güttel 2020).

A well-trained ANN can emulate, forecast or even create realistic results based on limited or incomplete nonlinear data (Schmidhuber 2014; Michelucci 2018). This data situation is given for the Japanese (or every other country's) siltation case, as demonstrated by Landwehr *et al.* (2020).

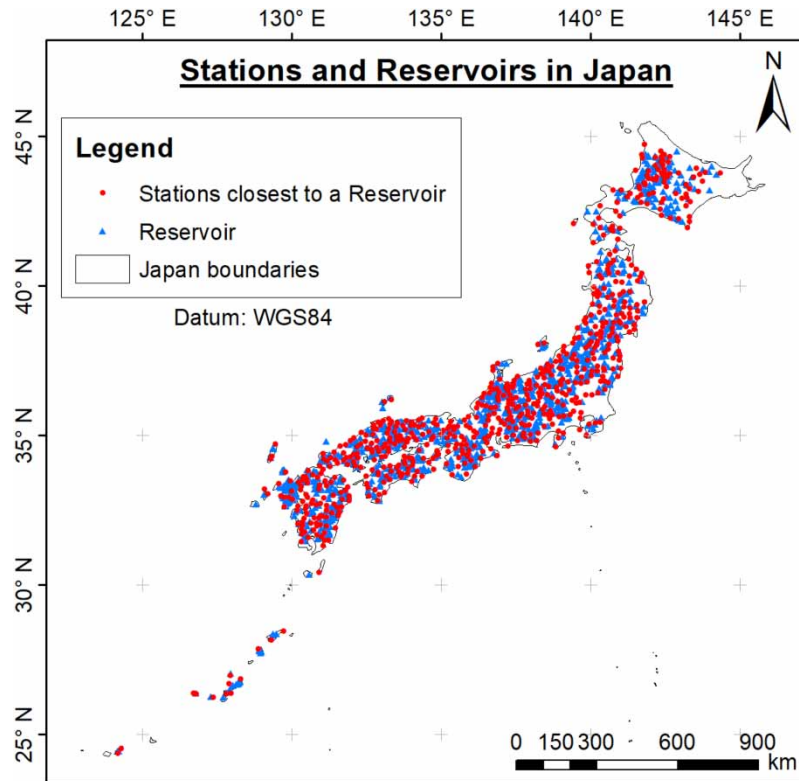


Figure 5 | Locations of reservoirs and precipitation stations in the data set.

An ANN processes input data in various steps via simple to complex logical and calculation operations. The ANNs are structured in layers. A single operation within that layer is called an artificial neuron. The artificial neurons are (just like the biological originals) highly interconnected.

In the case that every cell of a layer is connected with every cell of the next layer, that layer is called *fully connected*. Each neuron consists of adaptable weights which are the main lever for the flexibility of neural networks. An example of a simple ANN is displayed in Figure 6.

During a training phase, input data are processed, and the known R of the input data is provided. The result R' , which is produced by the ANN, is compared to R . Their difference is called the loss. Via a loss function, all weights of the network are adapted with the objective of producing an R' that is close to R . This is called backpropagation. All data are processed by the network in several iterations, the number of which is determined by the batch size. All data are referred to the network in multiple epochs, namely, multiple times. At the end of each epoch, a validation phase is conducted, in which the ANN capability is evaluated without the known result. This validation loss is a crucial measure of the progress of a network.

Afterward, a testing phase is conducted on data that are completely unknown to the network. Again, the result R is hidden from the network. If the ANN is sufficiently flexible to create R_{test} that is close to R_{test} , it is regarded as successful. In this case the ANN is capable of showing that it does not merely copy the data pattern, but instead learns inherent data characteristics by the selection of suitable weights (Schmidhuber 2014; Michelucci 2018). As this is exactly the objective of the present study (the detection of data characteristics and extraction of their influence from a vast data set, as illustrated in Figure 3), ANNs are an ideal tool.

2.2.1. Gated recurrent units

For instance, GRUs are utilized to preserve information from the second word/time step of a phrase/time series that influences the fifth word/time step of another phrase/time series. Whereas basic ANNs do not provide a memory function for preserving the necessary information, basic RNNs suffer from the vanishing gradient problem, where initial weights are no longer altered and no learning occurs, thereby leading to an *incorrect* memory. The GRU structure solves this problem

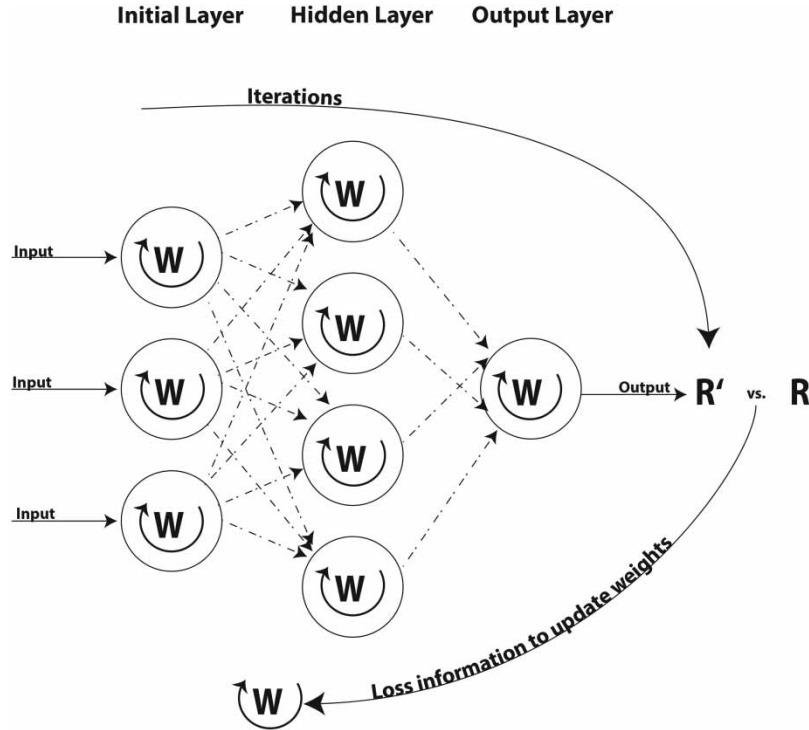


Figure 6 | Schematic diagram of a simple neuronal network with three layers and no special neurons, such as GRUs. The ANN produces a result R' via its weighted neurons during training that is expected to mimic (with a limited variance) the known result R . The loss function is used to constantly update and improve the weights. W serves as an abbreviation for all weights between the neurons.

with the update gatekeeper function Z_t from the following equation (Chung *et al.* 2014; Zhang *et al.* 2020):

$$Z_t = \sigma(X_t * W_{xz} + H_{t-1} * W_{hz} + b_z) \quad (1)$$

where Z_t is the updated gatekeeper function, X_t is the new information vector in time step t , H_{t-1} is the output information from the previous time step, W is the neuronal weight factor, b is the bias, xy is the affiliation of bias or direction of weight factor and σ is the sigmoid activation function.

It adds the new information X_t with the output H_{t-1} from the previous step. Consistent with the process of ANNs, the information is weighted by the iteratively updated weights W . The σ function scales the result to between 0 and 1. Z_t can be used as a gatekeeper that decides if and to what extent information is processed. Z_t 's gatekeeper functionality is applied in Equation (2):

$$H_t = Z_t \odot H_{t-1} + (1 - Z_t) \odot \tilde{H}_t \quad (2)$$

where \tilde{H}_t is the candidate output information, H_t is the GRU output information, \odot is the Hadamard (elementwise) product operator and Z_t decides the degrees of influence of both \tilde{H}_t and H_t on the output. It is also capable of turning off each one's influence completely. Thus, Z_t preserves old information when necessary or replaces or modifies it with new information from \tilde{H}_t . As the operation is performed elementwise via the \odot operator, the gatekeeper function selectively decides to update parts of the information. \tilde{H}_t is defined as follows:

$$\tilde{H}_t = \tanh(X_t * W_{xh} + R_t \odot H_{t-1} W_{hh} + b_h) \quad (3)$$

where R_t is the relevant gatekeeper function.

The added new and old information is scaled between -1 and 1 via \tanh . The information is gated via R_t which decides whether old information should influence the candidate output. R_t is displayed below and operates essentially like Equation (5):

$$R_t = \sigma(X_t * W_{xr} + H_{t-1} * W_{hr} + b_r) \quad (4)$$

The complete process is illustrated in Figure 7. The most important step is the updated gatekeeper mechanism, which enables selective propagation, updating and discarding of information.

This is a deciding key feature for the present study. It enables the overall network to learn and apply the characteristics of the occasionally notified event categories in training and testing. Therefore, GRU is an ideal tool for surveying the effects of event categories and realizing the desired task frame, as illustrated in Figure 3.

2.3. Data preparation

Precipitation measurement stations and reservoir locations were not concordant (see Figure 5). The Pythagorean theorem was applied to match each reservoir with its closest precipitation station (see Figure 5).

The data for reservoirs and precipitation is extensive. However, sedimentation or precipitation data was not continuous, complete or accurate in all cases. This was due to gaps in measurement, typing errors of operators or the simple fact that stations or reservoirs were just inaugurated after the year 2000.

Random forest imputation based on Stekhoven & Bühlmann (2011) was utilized to fill the gaps smoothly with the objective of imitating natural conditions that were derived from available data. To realize this objective, the algorithm utilizes the Gini impurity, which measures the likelihood of randomly chosen values fitting into a specified data distribution. The algorithm uses a likelihood divide that is based on the Gini impurity (it branches into two directions) and continues with branching until it reaches a level where it cannot further reduce the Gini impurity. Thus, it forms a tree. This methodology is applied various times with randomized values to create a random forest. The values with the lowest Gini impurity are chosen to fill the data gaps (Koehrsen 2018).

For unavailable data, an out-of-bound dummy value was chosen. In the case of nonconcordance, e.g., when siltated active and inactive storage would not sum to complete siltation, semiautomated search correction (a simple visual basic algorithm that adjusts values and, in case of inconsistencies, leaves the final verdict to humans and, therefore, is semiautomated) was applied.

Operators of the 1,224 reservoirs left notifications regarding special ISMTs and (natural) events. Those notifications were neither normed nor streamlined and were left to the respective employee's intention, interpretation and preferred Japanese syntax and semantics. Consequently, there was no unique pattern that could be interpreted by a neural network. Hence, the

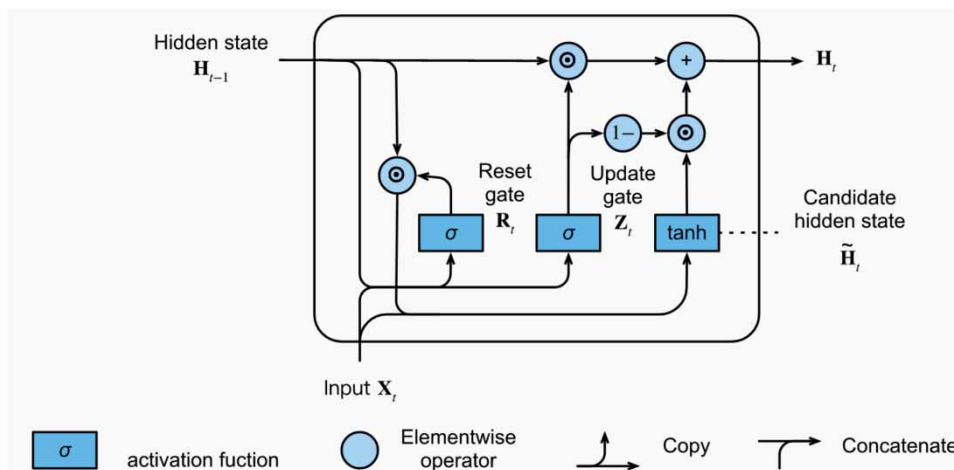


Figure 7 | Representation of operational processes of a GRU according to Zhang et al. (2020).

notifications were manually interpreted and unified into 14 overarching categories of events and ISMTs that were most common or frequent. They are listed in Table 1.

The categorized ISMTs and events were – similar to other categorical values – labeled and enumerated, which rendered them interpretable for machine learning.

ANN performance depends on the availability of vast data input with substantial data variety to avoid overfitting (Allamy 2014; Zhang *et al.* 2018). This is also valid for time-series surveys (Landwehr *et al.* 2020). The larger the dataset is, the higher the probability of an ANN adapting to unique data patterns and, thus, detecting the influence of the aforementioned event categories. Hence, it is necessary to artificially augment the data for two main reasons:

1. *To increase the absolute number of events with reduced appearance:* This helps the ANN detect rare events and learn from them. In the original dataset, various events are rare (or very rare); therefore, augmentation is mandatory.
2. *To increase the data variety:* Variety (e.g., event occurrence or siltation rise) helps the ANN not merely copy data patterns of the training set but also learn its inherent information. When unknown (test) data are faced, data variety (produced by data augmentation) helps the ANN flexibly adapt.

Thus, the available data were augmented using the TSAUG (time-series segmentation) algorithm based on Wen & Keyes (2019). It was selected because it enables the use of semi-intelligent augmentation techniques. The algorithm connects events to corresponding time-series patterns and induces semi-randomized varieties to artificially created copies of the original time-series compound.

The following techniques were used: 10–35% drift with 80% probability of occurrence (Um *et al.* 2017), a Blackman-Harris-based window function with 80% probability of occurrence (Nuttall 1981; Smith 2011) and randomized time warping (Um *et al.* 2017). These techniques shifted the occurrence of events and induced a variety of slopes and patterns. Each of the original 1,224 time series was augmented 10 times, thereby creating a data set of 13,464 reservoirs (1,224 of the original) with 18 time steps.

Each reservoir's 18 time step series were sequenced with a rolling-window algorithm that was used in forward projection to provide a more effective memory function of n consecutive years (de Meer Pardo 2019). The rolling window always uses the last predicted value and n of its predecessors. The initial prediction is determined by the original data values. n was chosen to be 4 in the present paper, which was determined by the lower loss and higher accuracy values of the subsequent ANN.

Finally, the data were separated into randomly organized test (10% of the data), training (81% of the data) and validation (9% of the data) sets according to the requirements of Section 2.2, which was carried out via ordinary numpy shuffle

Table 1 | Event categories from MLIT data alongside the prevailing technical terms of the Japanese original

Event Categories	イベント項目
Bypass	バイパス
Upstream Dam Installation	上流ダム設置
Unknown	不明
Flood/Taifun	出水や台風
Earthquake	地震
Sediment Relocation	堆砂移動
Dredging	屈出
New Observation/Calculation Methodology	新たな観測や計算方法
Alteration of Dam Volume/Height	貯水量の改変やダムの嵩上げ
Termination of the Reservoir	施設廃止
Spilling	放流
Drought	渇水
Measurement Error	測量誤差
Management Change	経営変化

commands. The size of the respective sets follows standard recommendations (Michelucci 2018). The sets were normalized via removing the mean and scaling to unit variance utilizing the subsequent equation:

$$z = \frac{x - \mu}{s} \quad (5)$$

where z is the standardized value, x is the sample from data set, μ is the mean of the data set and s is the standard deviation of the data set.

2.3.1. Target value

The target value that will help retrace the effects of events or ISMT on reservoir siltation must serve as a comparative value across all types and sizes of reservoirs. Thus, it must be relevant and capable of serving as a reference for reservoirs that are not part of the data set, namely, the target value must form a reference class for the generation of an outside view (Kahneman 2011; Flyvbjerg 2016). In contrast to optimistic bias detection (the classical use case of the outside view) (Landwehr *et al.* 2020), the effects of ISMT and events on siltation are paramount in the present paper.

Hence, a value that is based on a high variance is not recommended (such as inactive volume siltation objectives (Landwehr *et al.* 2020)). Rather, a stable relative value is needed. Therefore, siltation with respect to the whole reservoir volume is chosen as a comparative parameter for the formation of the desired reference class. To determine whether an event category indeed reduces or enforces siltation across a high variety of reservoirs, the impact of the event category must be detached from other key data. This is a highly complex mass data analysis process for which ANNs are suitable tools.

2.4. Applied GRU network and its hyperparameters

As described in Section 2.2.1, GRU layers proved to be a highly promising tool for time-series prediction data characteristic detection. They are even a valid alternative to the often-used LSTM units. In various cases, GRUs yielded better results due to their less complex operational structure (Che *et al.* 2016; Gao & Glowacka 2016; Wielgosz *et al.* 2017).

The ANN that was designed for this study consists of *bidirectional* GRU layers for preserving both past and predicted features of the LSTM. At this stage, underfitting (Allamy 2014) and overfitting (Allamy 2014; Zhang *et al.* 2018) issues are addressed with two dropout layers (Park & Kwak 2017). This is followed by a simple *time-distributed* fully connected layer. This serves to analyze each of the GRU outputs in time-series order. The output is followed by a flattening layer that reduces the dimension, which is necessary for generating the desired target value output. The network is concluded by two fully connected layers, which generate the output for the target value. The structure is illustrated in Figure 8. As the dropout and flattening layers do not consist of their own parameters, they are displayed in an in-between style.

ANNs are subject to hyperparameters that regulate them (see Table 2). These hyperparameters have a tremendous influence on the ANN performance (Nisbet *et al.* 2017; Michelucci 2018). For ANN design, setting of the hyperparameters is one of the most challenging steps (Claesen & Moor 2015).

To select the optimal operation environment for this study, the best-performing hyperparameters were chosen using an iterative cross-validation hyperband random search process (Jiang & Chen 2016; Li *et al.* 2018). In three attempts, the root-mean-square error (RMSE), valorization accuracy (Val-Acc) and Kullback–Leibler divergence (KLD) (Clim *et al.* 2018) were utilized as statistical guidance for the iterative optimization process. The promising median absolute deviation (MAD) was not used due to technical limitations (Gorard 2013; Landwehr *et al.* 2020).

While the Val-Acc and KLD optimizations did produce underfitted and, thus, not useful results, RMSE optimization produced hyperparameters that led to a promising low validation loss and, hence, a satisfactory fit for R' (see Section 2.2). The obtained hyperparameter values are listed in Table 2. They led to a network with 4,148,271 trainable values that were expected to be capable of deriving the desired information according to Figure 3.



Figure 8 | A simplified version of the developed network. FC, fully connected.

Table 2 | Overview of cross-validated hyperband random-searched hyperparameters

Hyperparameter	
1st GRU-layer size	542
2nd GRU-layer size	242
Time-distributed-layer size	242
Fully connected-layer size	100
1st GRU-layer activation	ReLU
2nd GRU-layer activation	ReLU
Time-distributed-layer activation	ReLU
Fully connected-layer activation	ReLU
1st dropout-layer percentage	0.45
2nd dropout-layer percentage	0.35
Optimizer	Nadam
Learning rate	0.000443
Loss	MSE
Batch size	242
Epochs	61

ReLU, rectified linear unit (Nwankpa *et al.* 2018); Nadam, Nesterov-accelerated adaptive moment estimation (Ruder 2017); MSE, mean squared error.

2.5. Metrics for GRU and result performance

The validation loss was selected as a metric for evaluating the training optimization. The ANN uses a loss function to measure and subsequently reduce the difference between the produced prediction and the already known results (Wang *et al.* 2018), as indicated in Section 2.2. The validation loss is calculated from the separated 9% of the data, which is provided to the network in each iteration. The lower the validation loss is, the better the network adapts to unknown data such as the test data. A perfect zero-loss match is, however, undesirable, as it would indicate overfitting on the validation data set (Allamy 2014; Zhang *et al.* 2018), thereby demonstrating that the internal training data variance is likely not sufficient for the creation of a highly adaptive network.

Hindcast emulations for each original or artificially produced reservoir data item of the test set are expected as results. With the hindcast approach, known data are treated as new data for the prediction of future data – with the real result used as a measure for control. In the case of a satisfactory match between the emulation and the original data, the general capability of emulating completely unknown data or completing forecasting tasks is demonstrated. Alongside the validation loss, the hindcast match comparison evaluates the general functionality of the trained neural network.

To examine the effects of the event categories, a semi-randomly selected variance of the original reservoir data was modified and retested. The only criteria were that selected decades of reservoir erection and all volume sizes of reservoirs were included in each possible combination within the selection; therefore, it was semirandom.

The data modification was restricted solely to the event categories. For each reservoir, the original event annotations were deleted, and an ‘injection’ of each event from the 14 event categories was undertaken for the period between 2007 and 2011. The neural network, which was trained on the original training data, subsequently executed an emulation for each of the modified reservoir data sets, namely, 15 (14 categories plus the original data) times 19 (five size categories times four selected decades minus one nonexistent combination) emulations were conducted. The process is illustrated in Figure 9.

The results were compared to the original data via a simple absolute median of the deviations. The median was selected because it reduces – in contrast to the average – the influence of outlier values on the statistical outcome.

Each of the event categories represents a unique distribution across the reservoirs. As the modified and original emulation data distributions reflect connected but not normally distributed data distributions, a Wilcoxon signed-rank test was conducted (Wilcoxon 1945) between the original distributions and those for each event category.

In case of significance, the null hypothesis that the modified and original data are of the same distribution, namely, that the event categories have no effect, would be rejected. The simple average would reflect the relative impact of each event category

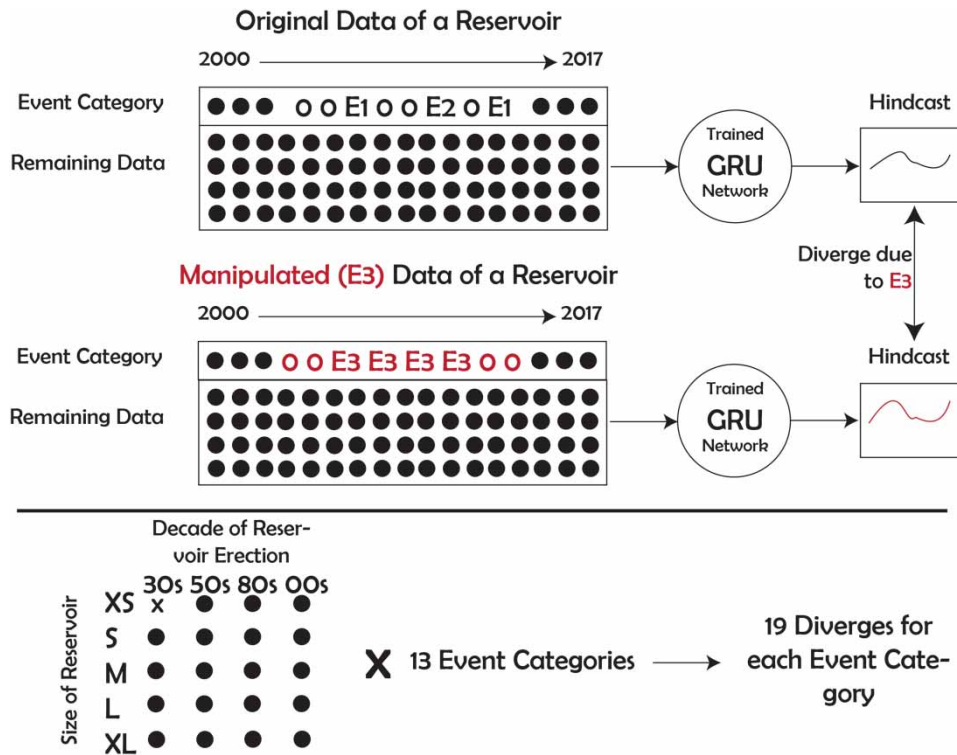


Figure 9 | Manipulation of the event categories causes the trained network to produce different hindcasts. The difference reflects the impact on siltation that the trained network has learned for the corresponding effect.

compared to the original distribution. This would imply that the neural network could actively derive and learn from the data whether each event category has an individual effect. Thus, interference by other side effects (other data) could be excluded. This concludes the methodological process of the present paper, as illustrated in Figure 4.

3. RESULTS

The data categories that were used for the training and testing of the network were event categories, reservoir size, precipitation for each month, year and initial sedimentation to the whole volume.

Variations (e.g., for batch size or epochs) of the hyperparameter optimal values from Section 2.4 were evaluated to obtain a better validation loss result. Nevertheless, the originally produced hyperband random-searched values proved to produce one of the best results with a validation loss of 0.3724 after 61 epochs. Better results from other configurations were insignificantly better, which is why the hyperband-optimized configuration was chosen for further analysis.

The network is designed to produce 15 emulations for each reservoir, namely, 15 hindcasts. With each new reservoir from the data, the network starts a new 15-step emulation turn. To display them more clearly, the emulations were structured linearly one after each other, as shown in Figure 10.

A hindcast match comparison and the validation loss demonstrated that the GRU network can predict convincingly on completely unknown reservoirs with extremely different configurations. It has learned to derive complex information from an extremely reduced essential data pool.

The subsequent objective was to disentangle the information and identify the influence of the network that was learned from the data for the event categories. This was conducted according to Section 2 and Figure 3.

Examples of the randomly selected reservoir cases and the respective influence of the ‘event category injections’ are presented in Figure 11. It is emphasized that Figure 11 represents an illustrative selection from the selection that followed the methodology explained in Section 2.5 and Figure 9. As the circumstances of each reservoir are highly individual, As the circumstances of each reservoir are highly individual, so are the ‘injection’ effects for each reservoir sedimentation emulation. They are ‘small pictures’.

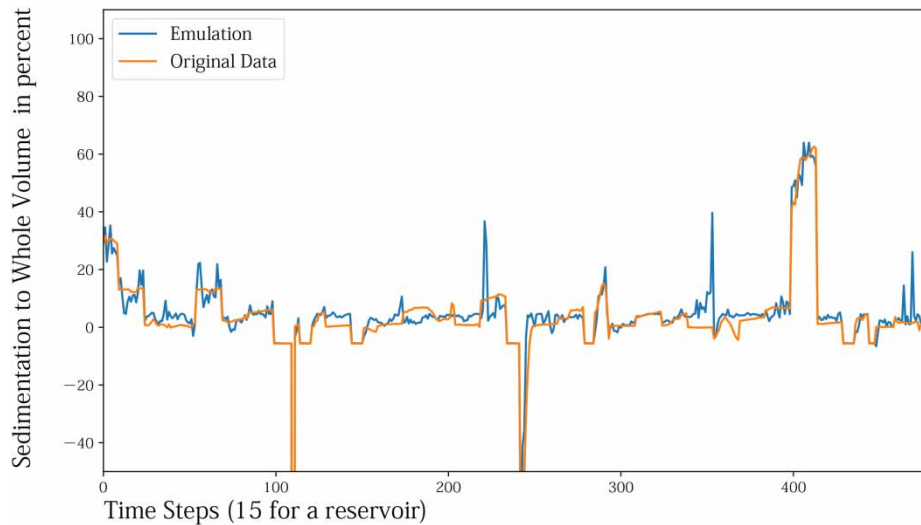


Figure 10 | Emulations of sedimentation to whole volume for 32 reservoirs in a randomized order. For every 15 time steps, a new reservoir is represented.

For example, the effect of ‘Floods/Typhoons’ on the rather recently inaugurated *Urayama dam* (浦山ダム) in Saitama is judged to be far more beneficial than in the three other and older dams, as presented in Figure 11. A reason for this might be ‘sluicing’, which might use more on novel dams. ‘Sluicing’ is debated in detail in Section 4.3.

The ‘big picture’ and the event category effect are just to be revealed with the overall statistical analysis and evaluation of all ‘injection’ cases below.

The effects of the event category ‘injections’ differ in terms of both extent and shape among the types of reservoirs. According to Section 2.5 and Figure 9, the event manipulation was established between 2007 and 2011. In concordance, no deviation effects from the original hindcast are visible before time step four in Figure 11. However, the influence of events continues to affect the hindcast for almost all subsequent time steps.

The effects are identified using the simple absolute median of the deviation from the original data, as described in Section 2.5. The null hypothesis is rejected for all 14 ‘injections’ with very high confidence in most cases, as the *Z*-values ranged far outside the quantile values (± 1.96 for 95% confidence).

Table 3 presents the effect that the GRU network has learned from the data for each event category with the methodology that is described in Section 2.5. The lower the absolute median deviation is, the larger the reduction in siltation per whole volume. The values are specified as percentages in comparison to the originally assigned event categories. For example, in the case of *dredging*, the median for the 19 emulation cases was 0.316% lower than the original hindcast of the GRU emulation. For reservoirs like the ones of the *Tsuruta Dam* (鶴田ダム) or the *Honna Dam* (本名ダム) from Figure 2, which both have roughly 8,000–9,000 m of sedimentation in 2017, this would mean 25,000 m less than sediment, which is roughly the plenary hall of the German Bundestag (Albers 1999). Thus, the median reduction via *dredging* is considerable.

The key message of Table 3 does, however, not reside within their absolute values but in the relation of the values to each other. Hence, according to the GRU network, *dredging* has a 29% higher impact on the reduction in siltation per whole volume than *spilling* (see column RM of Table 3).

4. DISCUSSION

The results provide compelling insight into the sedimentation management effect, as this is – to the best of the authors’ knowledge – one of the few, if not the only, mass data studies of siltation. However, the results are not entirely intuitive and, thus, are subject to discussion. In addition, the ANNs and the data input are worth discussing in order to better understand the current and optimize future research.

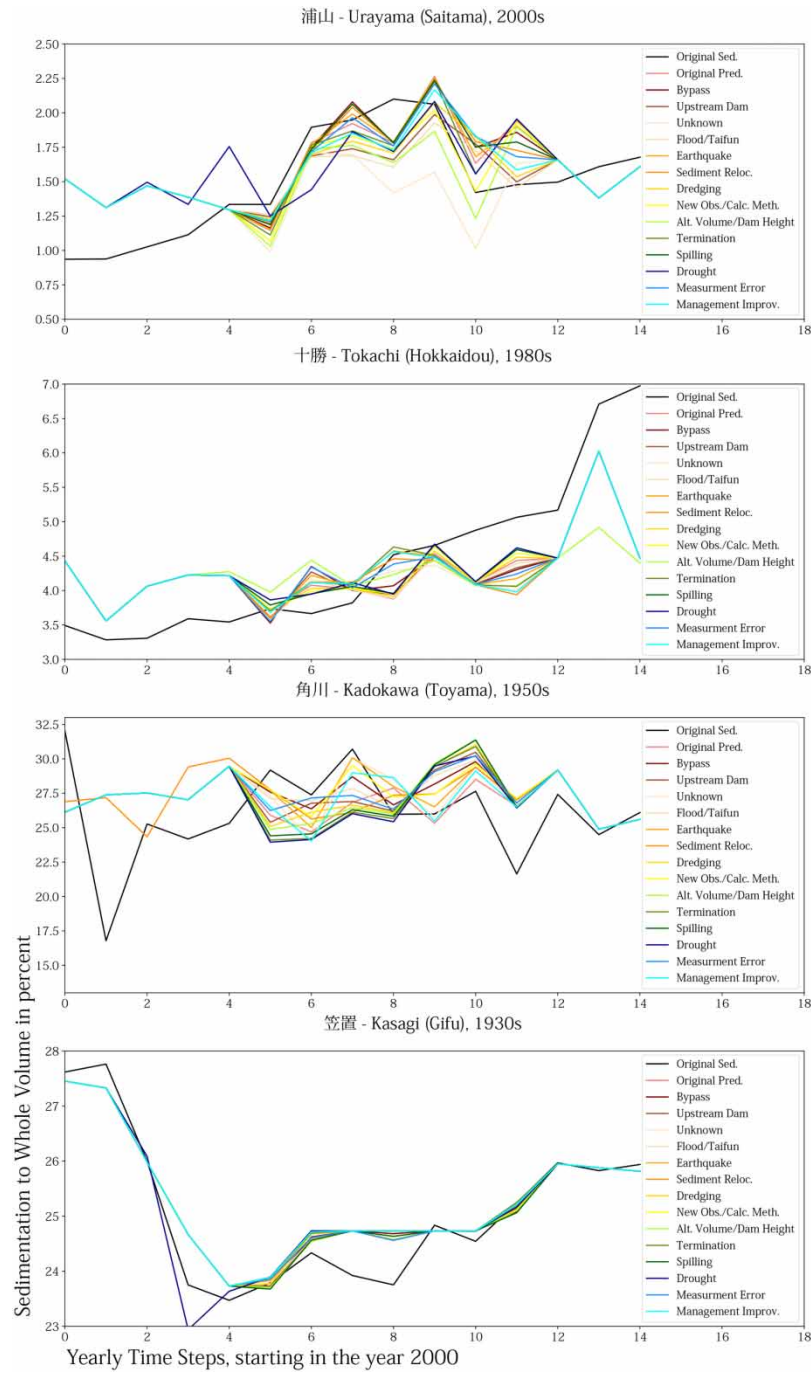


Figure 11 | Examples of the randomly selected reservoir cases: comparison of the originals with the 14 'event category injections'.

4.1. Event categories

Table 3 presents the efficiency of the GRU network for each event category. Before debating the implications of the results for various categories, it is important to closely examine the underlying data of the events.

4.1.1. Interpretation conflicts

The events are not a homogeneous mass. The MLIT received the event annotations from the operators.

Table 3 | Deviations among the semirandomly selected and event-manipulated reservoirs

AMD	RM	Event category	Wilcoxon Z-value
-0.31613	100.00000	Dredging	-5.68605
-0.30385	96.70048	Measurement error	-10.73597
-0.21024	71.54733	Spilling	-1.75157
-0.20797	70.93692	Sediment relocation	-8.60382
-0.14126	53.01029	Alteration of dam Volume/height	-5.19799
-0.13491	51.30526	Bypass	-6.58629
-0.13016	50.02848	Flood/typhoon	-1.92266
-0.07428	35.01180	Upstream dam Installation	-2.99249
-0.06804	33.33494	Management change	-5.92457
-0.02223	21.02507	Unknown	-1.94327
-0.01540	19.19086	Drought	-3.54715
-0.00069	15.23878	Termination of the reservoir	-2.16655
0.02835	7.43600	New Observation/calculation method	-1.89072
0.05602	0.00000	Eearthquake	-1.94858

AMD, absolute median deviation for sedimentation to whole volume in percent; RM, relative median, i.e., ranking of the lowest AMD (100%) compared to the highest AMD (0%); Wilcoxon Z-Value, approximation for critical values of the Wilcoxon signed-rank test to the normal deviation for $n > 20$ test values (Gibbons & Chakraborti 2011).

The MLIT has improved the format and standardization of siltation data over the years. The event annotations, however, are only vaguely standardized.

This has the advantage that operators can individually describe the highly diverse characteristics of reservoirs and events within Japan. However, the substantial variance of the descriptions from the operators often leaves excessive room for interpretation.

For example, the 2010 event notification 葦繁茂によりダム内への土砂流入をせき止めた – *Reed growth prevented sediment from flowing into the dam* of the Amagime (天君) reservoir in the Kumamoto Prefecture is the source of the following interpretation problems:

1. The notification could be incorporated into a larger category, namely, *Landscape Management*. Landscape management for sediment entry reduction is a hot topic for reservoirs (Sumi & Kantoush 2011; Kondolf *et al.* 2014). However, it is unclear whether other operators mention landscape management practices when they carry them out. In fact, the event notifications regarding landscape management were scant for the whole data set. It can thus be doubted that every landscape management practice was notified.
2. The reaction to the notification remained unclear. Was the reed cut? Was more reed planted? Certainly, there was a reaction, but its nature remains unknown.

In the end, the Amagime notification was included in the category *Management Change*. This is correct, as a management change was made. Nevertheless, as the true nature of the change is unclear, the probability that the effect of the Amagime measure conflicts with other entries of the *Management Change* event category is high. This might explain the modest effect that the GRU network derived for the *Management Change* category. *Management Change* might be much more efficient than indicated by the present paper, but event interpretation conflicts cause internal ambiguity.

Those types of interpretation conflicts did not prevail for the data set, but they were also not rare among the $1,224 \times 18 = 22,032$ possible event entries. Thus, the event categories are not free from impurities. The event reporting and its implications could be better, but the MLIT data in their current form already provide unique and valuable pieces of information.

4.2. Restricted occurrence

Various events suffered from significant application restrictions. A paramount example of this is the category *bypass*, which only has two occurrences within the data set. A *bypass* is presumed to be a highly impactful sediment management strategy

(Boes *et al.* 2014; Auel 2018). However, it is subject to tremendous application costs due to its complex construction characteristics (Morris 2020). This limits the application of this strategy drastically, as the data set demonstrates.

Due to the uniqueness and promising outlook of the technology for future applications (Morris 2020), it was decided to keep the *bypass* as a highly unique event category. The limitation effect was reduced by the application of the TSAUG methodology (Section 2.3). Disproportionate augmentation of the *bypass* cases was also considered, but it was decided that the base data restriction would contort the actual results too severely in this case. Nevertheless, it is clear that the base data restriction influenced the interpretation of the GRU network. Thus, the AMD and RM values are highly indicative of the *bypass* class.

The *bypass* case is an extreme example. Most of the other categories are founded on substantially more base data (several dozen to hundreds of event entries).

4.3. Spilling, sluicing, floods and typhoons

Spilling (流出–Ryuushutsu), which is also referred to as flushing, is among the most common sediment management techniques (Kantoush & Sumi 2017; Morris 2020). It is conducted by elevating the hydraulic scour. Although the definitions partially overlap, it should not be confused with sluicing, which typically involves the utilization of naturally occurring floods.

Interestingly, *Spilling* is considered to have substantially less impact than dredging, which is also often applied and which the GRU network found to be most efficient in sediment reduction.

Sluicing, in contrast, was not assigned its own category. This is due to two reasons:

1. It is sometimes treated homologously to *Spilling* by operators.
2. Flood events are used for sluicing, but depending on the type and fierceness of the flood (or typhoon) and the type of reservoir, sediment reduction or successful application of the sluicing technique is not always guaranteed. Thus, after flood event notifications, a sediment rise or decrease is observed within the data set. Nevertheless, it is impossible to distinguish flood events with sluicing from those without sluicing, as the information that is provided by the operator is restricted.

Therefore, the category was named Flood/Typhoon. Its data are not completely consistent. Nevertheless, it insinuates that Japanese operators are prepared to use (extreme) flood/typhoon events for sluicing since the overall GRU result demonstrates a reduction in sedimentation to the whole volume due to floods/typhoons.

4.4. Measurement errors, sediment relocations and new observation/calculation methods

Events need not always be intended or be of *vis major*. Frequent errors or unforeseen developments due to technological improvement impact sediment statistics. Three event categories are exemplary:

1. *Measurement error*: This is one of the most frequently occurring event entries. Hence, even in highly organized countries such as Japan, absolute certainty regarding data is not guaranteed. Moreover, it demonstrates that Japanese authorities and operators are willing to admit errors so that correct data can be obtained and higher operational sustainability realized. Interestingly, new measurements frequently insinuate overestimation errors in prior measurements.
2. *Sediment relocation*: Notification of this event was frequent. This is assumed to be due to the occurrence of floods/typhoons or the use of new observation/calculation methods. Additionally, it is possible that 堆砂移動 – *Sediment Relocation* is used as a filler declaration when the real reasons are unknown.
3. *New observation/calculation methods*: This event occurs frequently when advancing to more sophisticated bathymetric techniques (Balan *et al.* 2013; Kantoush & Sumi 2017; Adebayo Olubukola *et al.* 2020). Measured values can significantly deviate from previous values.

4.5. Data input

The data were gathered under the objective of universal applicability: the objective was to obtain a result of high significance without the necessity of assembling data that are not available in many countries due to restrictive governance structures (Pahl-Wostl 2015). Thus, more accurate results are likely to be obtained with more data input.

However, the present paper demonstrates that highly promising results can already be produced with only a handful of influential factors.

The exact relation between most of the data and event categories remains a subject for further investigation (e.g., whether there is a satisfactory correlation between flood/typhoon events and the precipitation data and its impact on sedimentation).

4.6. Networks

There are many approaches that might lead to advanced ANNs that can yield more accurate results. LSTM layers are a possible alternative for GRU layers, and convolution and average pooling layers are also interesting approaches for information extraction (Tang *et al.* 2020). The same might be true for a more extensive use of time distribution layers.

Another opportunity to generate more realistic artificial data is assumed to lie within the use of generative adversarial networks, which have already produced convincing results for artificial image generation (Han *et al.* 2019; Islam & Zhang 2020).

The ANN opportunities are vast and were partially evaluated in this study. The eventually utilized GRU network proved to be the most reliable network that yielded the best results with the resources that were available at the time of the study.

5. CONCLUSIONS AND OUTLOOK

The presented methodology has high potential for scenario testing of management strategies. The impacts of highly complex technologies on sedimentation compartment can be tested with moderate effort and data demand to realize hindcasts with satisfactory performance. The foundation of real data corresponds to the outside view and reference class that are demanded by Kahneman (2011) and Flyvbjerg (2016) for grand infrastructures.

Reservoirs are objects of high variance for multiple reasons, which range from climate and purpose to technology. Consequently, the individual characteristics of reservoirs play a key role in classical sediment development analysis. However, the ANN methodology enables both the derivation of conclusions from the big general picture of several hundred to thousands of reservoirs and attribution to individual factors. This is the main advantage of the presented methodology. Engineering can benefit from this research directly as the mass data approach can reveal for which case certain anti-sedimentation strategies like spilling gates or bypass tunnels are more effective and would thus justify this high-cost investment. Hence, planning would gain another key tool to optimize existing and future dams based on results and experiences derived from well-trained neural networks. Nonetheless, the presented methodology cannot replace highly detailed individual reservoir analysis. The ANN methodology is a useful supplement.

In the special case of Japan, continuation and enhancement of the data set are paramount. More details for the event categories might be revealed if operators are approached directly (however, this is connected with a very high workload).

The potential power that governmental approaches possess is emphasized. Governmental and management structures and decisions may strongly influence sediment accumulation and, thus, the longevity of reservoirs. They incorporate the smart design and application of sediment technological countermeasures, such as spilling, dredging, sluicing or bypassing, that the results demonstrate to be efficient. They also incorporate adaptive measures, such as the modernization and realization of observation and control (measurement errors and new observation/calculation methods) and the formation of new management structures (management change). Overall, the results demonstrate that technological progress cannot exert its full effect if it is not supported and distributed by overarching governmental and management structures.

Further investigation of the data set with other approaches is necessary. Additional data classes are likely to provide further insight into various research questions, e.g., with respect to landscape management. Upcoming multivariate, multi-time-series approaches with various ANN methodologies will produce even more detailed findings. Especially the presumed opportunities of general adversarial networks (GAN) to mimic naturally existing data sets are emphasized. A breakthrough GAN study regarding data structures would facilitate data enhancement in a supposedly more natural manner.

Since the proposed approach is based on a reduced variety of data input, global applicability is assumed. Countries with scant data availability can benefit from either the methodology or pretrained networks.

ACKNOWLEDGEMENTS

This research was supported by the Environment Research and Technology Development Fund (No. JPMEERF20S11814) of the Environmental Restoration and Conservation Agency of Japan and JSPS KAKENHI under Grant No. JP21H01434.

DECLARATION OF INTEREST

I declare that this research is my own work except where there is clear acknowledgment and reference to the work of others. This research does not contain material that has already been used to any substantial extent for a comparable purpose.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

- Adebayo Olubukola, O., Ajani, S. & Akinwale, 2020 *Comparison of Different Methods for Bathymetric Survey and Sedimentation Evaluation of a Small Reservoir in Nigeria for Sustainable Management*. pp. 181–195.
- Adeogun, A. G., Sule, B. F. & Salami, A. W. 2018 *Cost effectiveness of sediment management strategies for mitigation of sedimentation at Jebba Hydropower reservoir, Nigeria*. *Journal of King Saud University – Engineering Sciences* **30** (2), 141–149.
- Albers, K.-J. 1999 *Besonderheiten der Raumluftechnischen Anlage des Plenarsaals im Umgebauten Reichstag Berlin*. Available from: https://www.hs-esslingen.de/fileadmin/media/Fakultaeten/gu/Forschungund/Labore/Veroeffentlichungen/Besonderheiten_der_raumluftechnischen_Anlagedes_Plenarsaals_im_umgebauten_ReichstagBerlin.pdf (accessed 19 January 2021).
- Allamy, H. 2014 *Methods to Avoid Over-Fitting and Under-Fitting in Supervised Machine Learning (Comparative Study)*.
- Annandale, G. 2013 *Quenching the Thirst: Sustainable Water Supply and Climate Change*.
- Auel, C. 2018 *Sediment Bypassing – A Sustainable and Eco-Friendly Strategy Against Reservoir Sedimentation*.
- Auel, C., Kantoush, S. & Sumi, T. 2016 *Positive Effects of Reservoir Sedimentation Management on Reservoir Life – Examples from Japan*.
- Balan, M., Das, C. D. A., Khandelwal, M. & Chaudhari, P. 2013 Review of various technologies for depth measurement in estimating reservoir sedimentation. *International Journal of Engineering Research and Technology* **2**, 223–228.
- Basson, G. R. 2009 Management of siltation in existing and new reservoirs. In: *23rd Congress of the Int. Commission on Large Dams CIGBICOLD*.
- Boes, R., Auel, C., Müller-Hagmann, M. & Albayrak, I. 2014 *Sediment Bypass Tunnels to Mitigate Reservoir Sedimentation and Restore Sediment Continuity*. pp. 221–228.
- Che, Z., Purushotham, S., Cho, K., Sontag, D. & Liu, Y. 2016 Recurrent neural networks for multivariate time series with missing values. *Scientific Reports* **8**, 6085.
- Chung, J., Gülçehre, Ç., Cho, K. & Bengio, Y. 2014 Empirical evaluation of gated recurrent neural networks on sequence modeling. *CoRR*, abs/1412.3555.
- Claesen, M. & Moor, B. D. 2015 *Hyperparameter Search in Machine Learning*.
- Clim, A., Zota, R. D. & Tinić, G. 2018 *The Kullback-Leibler divergence used in machine learning algorithms for health care applications and hypertension prediction: a literature review*. *Procedia Computer Science* **141**, 448–453. The 9th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN-2018)/The 8th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH-2018)/Affiliated Workshops.
- de Meer Pardo, F. 2019 *Enriching Financial Datasets with Generative Adversarial Networks*.
- de Vente, J., Poesen, J., Verstraeten, G., Govers, G., Vanmaercke, M., Rompaey, A. v., Arabkhedri, M. & Boix-Fayos, C. 2013 Predicting soil erosion and sediment yield at regional scales: where do we stand? *Earth-Science Reviews* **127**, 16–29.
- Dogmus, Ö. C. & Nielsen, J. 2019 *Is the hydropower boom actually taking place? a case study of a south east European country, Bosnia and Herzegovina*. *Renewable and Sustainable Energy Reviews* **110**, 278–289.
- Elsworth, S. & Güttel, S. 2020 *Time Series Forecasting Using LSTM Networks: A Symbolic Approach*.
- Flyvbjerg, B. 2016 How to get your business case right, by using Kahneman's outside view. *SSRN*.
- Fu, R., Zhang, Z. & Li, L. 2016 Using LSTM and GRU neural network methods for traffic flow prediction. In: *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. pp. 324–328.
- Gamboa, J. C. B. 2017 *Deep Learning for Time-Series Analysis*.
- Gao, Y. & Glowacka, D. 2016 *Deep Gate Recurrent Neural Network*.
- Ghimire, G. & DeVantier, B. 2016 *Sediment Modeling to Develop a Deposition Prediction Model at the Olmsted Locks and Dam Area*. pp. 410–420.
- Gibbons, J. D. & Chakraborti, S. 2011 *Nonparametric Statistical Inference*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 977–979.
- Gorard, S. 2013 The possible advantages of the mean absolute deviation 'effect' size. *Social Research Update* **65**, 1.
- Han, C., Rundo, L., Araki, R., Nagano, Y., Furukawa, Y., Mauri, G., Nakayama, H. & Hayashi, H. 2019 *Combining noise-to-image and image-to-image GANs: brain MR image augmentation for tumor detection*. *IEEE Access* **7**, 156966–156977.
- Hao, C.-F., Qiu, J. & Li, F.-F. 2017 *Methodology for analyzing and predicting the runoff and sediment into a reservoir*. *Water* **9** (6), 440–456.
- Haregeweyn, N., Tsunekawa, A., Tsubo, M., Meshesha, D., Poesen, J., Nyssen, J. & Deckers, J. 2013 Reservoir sedimentation and its mitigation strategies: a case study of the Ethiopian highlands. In *Hydro13 Proceedings*. p. 8.
- Islam, J. & Zhang, Y. 2020 *GAN-based synthetic brain pet image generation*. *Brain Informatics* **7**, 231–247.
- Jiang, P. & Chen, J. 2016 *Displacement prediction of landslide based on generalized regression neural networks with k-fold cross-validation*. *Neurocomputing* **198**, 40–47.
- Kahneman, D. 2011 *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York.
- Kantoush, S. & Sumi, T. 2010 *River Morphology and Sediment Management Strategies for Sustainable Reservoir in Japan and European Alps*. p. 53.

- Kantoush, S. A. & Sumi, T. 2017 *The Aging of Japan's Dams: Innovative Technologies for Improving Dams Water and Sediment Management, River Sedimentation*.
- Kantoush, S. A. & Sumi, T. 2019 *Paradigm Shift for Sediment Management*. Available from: <https://www.waterpowermagazine.com/features/featureparadigm-shift-for-sediment-7263873/> (accessed 23 December 2019).
- Koehrsen, W. 2018 *An Implementation and Explanation of the Random Forest in Python*. Available from: <https://towardsdatascience.com/an-implementation-and-explanation-of-the-random-forest-in-python-77bf308a9b76> (accessed 2 January 2021).
- Kondolf, G. M., Gao, Y., Annandale, G. W., Morris, G. L., Jiang, E., Zhang, J., Cao, Y., Carling, P., Fu, K., Guo, Q., Hotchkiss, R., Peteuil, C., Sumi, T., Wang, H.-W., Wang, Z. & Wei, Z. 2014 Sustainable sediment management in reservoirs and regulated rivers: experiences from five continents. *Earth's Future* **2**, 256–280.
- Landwehr, T., Kantoush, S. A., Pahl-Wostl, C., Sumi, T. & Irie, M. 2020 The effect of optimism bias and governmental action on siltation management within Japanese reservoirs surveyed via artificial neural network. *Big Earth Data* **4** (1), 68–89.
- Lee, H.-Y., Lin, Y.-T. & Chiu, Y.-J. 2006 Quantitative estimation of reservoir sedimentation from three typhoon events. *Journal of Hydrologic Engineering* **11** (4), 362–370.
- Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A. & Talwalkar, A. 2018 *Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization*.
- Michelucci, U. 2018 *Applied Deep Learning: A Case-Based Approach to Understanding Deep Neural Networks*.
- Morris, G. 2020 Classification of management alternatives to combat reservoir sedimentation. *Water* **12** (3), 861.
- Morris, G. L. & Fan, J. 1998 *Reservoir Sedimentation Handbook*, Vol. 1.01. McGraw-Hill Book Co, New York, NY.
- Nisbet, R., Miner, G. & Yale, K. 2017 *Handbook of Statistical Analysis and Data Mining Applications*, 2nd edn. Academic Press, Inc., Orlando, FL, USA.
- Nuttall, A. H. 1981 *Some Windows with Very Good Side-Lobe Behavior*.
- Nwankpa, C., Ijomah, W., Gachagan, A. & Marshall, S. 2018 *Activation Functions: Comparison of Trends in Practice and Research for Deep Learning*.
- Oehy, C. 2003 Effects of obstacles and jets on reservoir sedimentation due to turbidity currents, Ph.D. thesis, No. 15, Swiss Federal Institute of Technology, Lausanne, Switzerland. <https://infoscience.epfl.ch/record/116164>
- Omer, A. Y. A., Ali, A. S. A., Roelvink, J. A., Paron, P. & Crosato, A. 2015 Modelling of sedimentation processes inside Roseires Reservoir (Sudan). *Earth Surface Dynamics* **3**, 223–238.
- Pahl-Wostl, C. 2015 *Water Governance in the Face of Global Change: From Understanding to Transformation*.
- Pandey, A., Chaube, U., Mishra, S. & Kumar, D. 2016 Assessment of reservoir sedimentation using remote sensing and recommendations for desilting Patratu Reservoir, India. *Hydrological Sciences Journal* **61** (4), 711–718.
- Park, S. & Kwak, N. 2017 *Analysis on the Dropout Effect in Convolutional Neural Networks*. pp. 189–204.
- Petneházi, G. 2019 *Recurrent Neural Networks for Time Series Forecasting*.
- Pousse, E. & Latouche, D. 2018 *Transforming Our World: The 2030 Agenda for Sustainable Development*. Working Papers, Network of African Local Government Financing Institutions (RIAFICO).
- Ruder, S. 2017 *An Overview of Gradient Descent Optimization Algorithms*.
- Schleiss, A., De Cesare, G. & Althaus, J. J. 2010 Verlandung der Stauseen gefährdet die nachhaltige Nutzung der Wasserkraft. *Luft, Wasser, Energie*. <https://infoscience.epfl.ch/record/147714>.
- Schleiss, A., Franca, M., Juez, C. & De Cesare, G. 2016 Reservoir sedimentation. *Journal of Hydraulic Research* **54**, 1–20.
- Schmidhuber, J. 2014 Deep learning in neural networks: an overview. *Neural Networks* **61**, 85–117.
- Simoes, F. J. M. & Yang, C. T. 2006 *Sedimentation Modeling for Rivers and Reservoirs*. USGS, Reston, VA.
- Smith, J. O. 2011 *Spectral Audio Signal Processing*. Available from: <http://ccrma.stanford.edu/jos/sasp//ccrma.stanford.edu/~jos/sasp/>. online book, 2011 edition.
- Stähly, S., Franca, M., Robinson, C. & Schleiss, A. 2019 Sediment replenishment combined with an artificial flood improves river habitats downstream of a dam. *Scientific Reports* **9** (1), 5176–5184.
- Stekhoven, D. J. & Bühlmann, P. 2011 Missforest – non-parametric missing value imputation for mixed-type data. *Bioinformatics* **28** (1), 112–118.
- Sumi, T. 2015 Comprehensive reservoir sedimentation countermeasures in Japan. In: *International Workshop on Sediment Bypass Tunnels*.
- Sumi, T. & Kantoush, S. 2011 *Sediment Management Strategies for Sustainable Reservoir*. pp. 353–362.
- Tang, W., Long, G., Liu, L., Zhou, T., Jiang, J. & Blumenstein, M. 2020 *Rethinking 1D-CNN for Time Series Classification: A Stronger Baseline*.
- Um, T. T., Pfister, F. M., Pichler, D., Endo, S., Lang, M., Hirche, S., Fietzek, U. & Kulic, D. 2017 Data augmentation of wearable sensor data for Parkinson's disease monitoring using convolutional neural networks. In: *Proceedings of 19th ACM International Conference on Multimodal Interaction (ICMI'17)*, Glasgow. ACM.
- United Nations 2015 *Transforming Our World: The 2030 Agenda for Sustainable Development*. Working Papers, eSocialSciences.
- Vanmaercke, M., Kettner, A. J., Eeckhaut, M. V. D., Poesen, J., Mamaliga, A., Verstraeten, G., Rădoane, M., Obreja, F., Upton, P., Syvitski, J. P. & Govers, G. 2014 Moderate seismic activity affects contemporary sediment yields. *Progress in Physical Geography: Earth and Environment* **38** (2), 145–172.

- Velásquez-Castro, K., Ingol-Blanco, E., Pehovaz-Alvarez, R. & Cruzado-Blanco, C. 2016 *Assessment of Reservoir Sedimentation and Mitigation Measures: A Case Study of Palo Redondo Reservoir*.
- Wang, H.-W., Kondolf, M., Tullos, D. & Kuo, W.-C. 2018 *Sediment management in Taiwan's reservoirs and barriers to implementation*. *Water* **10** (8), 1034.
- Wen, T. & Keyes, R. 2019 *Time Series Anomaly Detection Using Convolutional Neural Networks and Transfer Learning*.
- Wielgosz, M., Skoczeń, A. & Mertik, M. 2017 *Recurrent Neural Networks for Anomaly Detection in the Post-Mortem Time Series of LHC Superconducting Magnets*.
- Wilcoxon, F. 1945 *Individual comparisons by ranking methods*. *Biometrics Bulletin* **1** (6), 80–83.
- Wild, T. B., Loucks, D. P., Annandale, G. W. & Kaini, P. 2016 *Maintaining sediment flows through hydropower dams in the Mekong river basin*. *Journal of Water Resources Planning and Management* **142** (1), 05015004.
- Yang, S.-q. 2013 *Why cannot sediment transport be accurately predicted*. In: *35th IAHR World Congress*. IAHR.
- Zarfl, C., Lumsdon, A. E., Berlekamp, J., Tydecks, L. & Tockner, K. 2015 *A global boom in hydropower dam construction*. *Aquatic Sciences* **77**, 1015–1621.
- Zeleeke, T., Moussa, A. M. & El-Manadely, M. S. 2013 *Prediction of sediment inflows to Angereb dam reservoir using the SRH-1D sediment transport model*. *Lakes & Reservoirs* **18**, 366–371.
- Zhang, Q., Prouty, C., Zimmerman, J. B. & Mihelcic, J. R. 2016 *More than target 6.3: a systems approach to rethinking sustainable development goals in a resource-scarce world*. *Engineering* **2** (4), 481–489.
- Zhang, C., Vinyals, O., Munos, R. & Bengio, S. 2018 *A Study on Overfitting in Deep Reinforcement Learning*.
- Zhang, A., Lipton, Z. C., Li, M. & Smola, A. J. 2020 *Dive Into Deep Learning*. Available from: <https://d2l.ai/chapter-recurrent-modern/gru.html> (accessed 23 December 2020).

First received 19 January 2021; accepted in revised form 28 November 2021. Available online 7 February 2022