The effect of optimism bias and governmental action on siltation management within Japanese reservoirs surveyed via artificial neural network

Tobias Landwehr, Sameh Ahmed Kantoush, Claudia Pahl-Wostl, Tetsuya Sumi & Mitsuteru Irie

Published online: 05 Feb 2020.

Submit your article to this journal

Article views: 1330

View related articles

View Crossmark data

Citing articles: 5 View citing articles

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

The effect of optimism bias and governmental action on siltation management within Japanese reservoirs surveyed via artificial neural network

Tobias Landwehr[a], Sameh Ahmed Kantoush[b], Claudia Pahl-Wostl[a], Tetsuya Sumi[b] and Mitsuteru Irie[c]

[a]Institute for Environmental Systems Research, Osnabrück University, Osnabrück, Germany; [b]Water Resources Research Center, Disaster Prevention Research Institute, Kyoto University, Kyoto, Japan; [c]Department of Civil and Environmental Engineering, Faculty of Engineering, Miyazaki University, Miyazaki, Japan

ABSTRACT

Reservoirs are 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. The present paper attempts to evaluate the global reservoir siltation problem with the optimism bias theorem introduced by Kahneman and Tversky and applied to infrastructural mega-projects by Flyvbjerg and Ansar using artificial neural networks (ANNs) algorithms for large Japanese reservoirs. Japan possesses suitable long-term data and a legal directive concerning the sediment capacity siltation duration, which serves as a valid guide to check whether, over the past 100 years, engineers, planners and managers were capable of judging the sediment input correctly. Various ANN models were established to emulate Japanese reservoir siltation behavior. The networks demonstrate that reservoirs in Japan suffer from optimism bias. In contrast to the law, the dead storage volume of an average dam is supposed to reach capacity after 52 years. This finding joins the overall observation that mega-projects generally and globally suffer from optimism bias. The emulations were subsequently screened for a presumed influence of governance actions, namely, indicating plus monitoring and the change in the market competition situation. While reservoir siltation appears to continue regardless of the level of competition in public procurement, monitoring directives appear to have a considerable impact on improved siltation management, which demonstrates that dedicated governance action can significantly strengthen the sustainable behavior of key infrastructure elements such as reservoirs.

1. Introduction

Recently, hydropower has experienced a tremendous global construction boom. Fueled by the desire to create more active storage (Annandale, 2013), the number of installed

dams is expected to more than double up to 2030 compared to 2010 (Zarfl, Lumsdon, Berlekamp, Tydecks, & Tockner, 2015). One of the reasons for this boom is siltation, which has led the gross reservoir volume per capita (Annandale, 2013) and the gross reservoir volume (Kantoush & Sumi, 2010) (as seen in Figure 1) shrink to 1970s levels despite the ever-growing number of reservoirs. This situation is a giant setback for sustainable long-term renewable energy generation and water supply security.

Notwithstanding manifold siltation predictions and elaborate simulation models (Ghimire & DeVantier, 2016; Hao, Qiu, & Li, 2017; Omer, Ali, Roelvink, Paron, & Crosato, 2015; Simoes & Yang, 2006; Zeleke, Moussa, & El-Manadely, 2013), the global development insinuates that planners, operators and practitioners struggle to predict both the sediment inflow and yield reservoirs correctly. Consequently, the implementation of suitable management strategies for siltation remains a challenge (de Vente et al., 2013; Kantoush & Sumi, 2017; Schleiss, Franca, Juez, & De Cesare, 2016; Yang, 2013), which prevails regardless of the respective environment, climate, society and technology level Annandale (2013); Basson (2009); Schleiss, De Cesare, and Althaus (2010). It appears that planners, operators, managers and constructors are overly optimistic, i.e. they are biased by the so-called optimism bias (Schleiss et al., 2016).

Said optimism allows responsible persons to underestimate risks while overestimating their own capacities. By doing so, they neglect the given data and experience of already installed projects of the same or similar kind (Ansar, Flyvbjerg, Budzier, & Lunn, 2014; Flyvbjerg, 2006, 2016). This theory was initially developed by Tversky and the Nobel prize-winning Kahneman (Kahneman & Tversky, 1979).

Siltation data are, however, globally scarce (Schleiss et al., 2016). Few authorities monitor sediment in a regular, overarching manner, except in Japan (Auel, Kantoush, & Sumi, 2016; Kondolf et al., 2014), which makes it an ideal country for data-driven experience analysis.

Because there is increasing pressure on reservoirs due to (unforeseen) siltation, retrospective amendment methods play a key role in sediment management (Annandale, 2013; Kantoush & Sumi, 2017; Schleiss et al., 2016). Said methods are manifold and often technical in nature; however, they are not generally applied (Kantoush & Sumi, 2017; Morris & Fan, 1998).
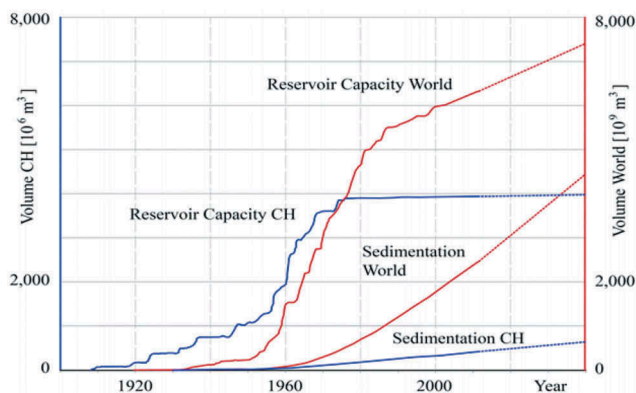
Figure 1. Assumed global (and Swiss) siltation development according to Kantoush and Sumi (2010).

On the other hand, governance measures exist. Similar to technological solutions, their use case application is also limited (Pahl-Wostl, 2015), which is indeed startling, as it is observed that governance measures can have a high impact within the water-energy-food nexus (WEF nexus) (Pahl-Wostl, 2015). Reservoirs, as a junction between energy production and water supply (for, e.g., crop production), belong to the WEF nexus. As one of its aims, this study seeks to find evidence that substantiates the effect of governance on sediment management.

Siltation is a non-linear time series process (Annandale, 2013). An emergent tool for non-linear data analysis and non-linear data emulation are artificial neural networks (ANNs), whose advantages include comparatively rapid applicability and a vast use case flexibility (Gamboa, 2017). Thus, they are this study's method of choice to detect evidence of optimism bias and governance effects.

The present paper seeks to identify evidence that substantiates the presumption of optimism bias behavior in sediment management and planning.

To verify whether optimism bias is indeed persistent in siltation prognosis and reservoir management, the outside view introduced by Kahneman and Tversky and data analysis are key features (Ansar et al., 2014; Kahneman & Tversky, 1979).

The subsequent methodologies will produce general results, i.e. this analysis does not account for every single dam in a highly specific manner – it is acknowledged that there are huge individual differences. Nevertheless, if the Japanese reservoirs were regarded as an entity, the produced results will give worthwhile insights about large scale reservoir sustainability.

2. Methodology

2.1. Selection of an optimism bias reference

Optimism bias detection requires a reference class (Ansar et al., 2014; Flyvbjerg, 2006; Kahneman, 2011) to establish a needed outside view(Kahneman, 2011; Kahneman & Tversky, 1979). More than often, (large scale) projects were established and carried out by the project team itself or by people related to the project judging on project data. This is the so-called inside view(Kahneman & Tversky, 1979). It is often quite sophisticated and accounts for the own resources, capacities and skills. However, what is unaccounted for are the so-called *unknown unknowns*, which are impossible to perceive if one resides inside of a system (Kahneman, 2011).

In doing so, large scale infrastructure projects are more than often affected by cost overshoot or by a project completion way beyond the initial schedule (Ansar et al., 2014; Flyvbjerg, 2016, 2006). An outside view is thus needed. It is provided by already completed projects of a similar class, i.e. a reference class is needed (Kahneman, 2011; Kahneman & Tversky, 1979). The reference class serves as a great indication, where projects are going to head in the future.

Translated to the case of Japanese reservoir siltation, the projects, i.e. reservoirs, that have already been installed and their sedimentation data are the very key to create a reference class. They are supposed to indicate quite neatly the overall siltation state of reservoirs after a certain period according to Flyvbjerg (2016); Kahneman and Tversky

(1979). To judge whether the outside view and the inside view diverge (i.e. whether an optimism bias exists), a siltation planning marker is needed.

Reservoirs have an installed siltation capacity, which is often congruent with its *Dead Storage Volume* (Kantoush & Sumi, 2017). This volume lies below the reservoir outlet and cannot be actively used for water management. However, it serves as siltation buffer to guarantee the full usage of active storage over a planned period of time. An insufficient siltation capacity is seen as a clear threat to operation security and a long reservoir lifetime (Kantoush & Sumi, 2017; Rakhmatullaev, 2010).

In this study, lifetime is defined as the termination of the initially intended operation purpose due to siltation; it does not necessarily equal a 100% siltated reservoir. By no means lifetime is equal to the ratio of siltation capacity to the mean annual sediment inflow. The siltation capacity situation is, however, a very good proxy to judge whether the intended operation lifetime or longevity will be kept in the future.

The time of operation for reservoirs must be estimated, designed and planned by managers, practitioners and engineers. A divergence from these plans is a clear indication of siltation underestimation. If siltation underestimation is systemic for reservoirs, it becomes a general bias, namely, an optimism bias.

For most of its reservoirs, Japan introduced (*Shouwa 32* -昭和 32 年) a directive in 1957 that demands one hundred years of minimum durability for the siltation capacity (*Kasen Hobou Gijutsu Kihon* – 河川破防技術基準 – Norm for relevant techniques against destructive river basin management, NTDM), according to Okano, Takanagi, Fujii, and Ando (2004).

The NTDM norm is the marker of this research. A general divergence of siltation in the sediment capacity volume in Japanese reservoirs from the legal siltation goal is seen as an indication of optimism bias in sediment prognosis and management.

2.2. Base data

The base data originated in 2015 (*Heisei 27* – 平成27年) and were provided by the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) and the Public Works Research Institute (PWRI). In total, 1006 dams are included (as shown in Figure 8). Dams of different purposes, types and sizes (small: <1 m m^3, medium: >1 m m^3, large: >10 m m^3; very large: >100 m m^3 of total storage volume [this definition is made for Japan and not internationally applicable]) are incorporated. Hence, the data reflect the general character of Japan's reservoir landscape. The data comprise the siltated volume, inactive storage capacity, and active storage capacity along with general data such as inauguration or river basin affiliation. Of these, 141 dam sets were not directly usable because they were incomplete. Stochastic measures were undertaken to utilize them (Figure 2).

As the data did not separate active and dead storage siltation, the assumption was made that the whole siltation occurred in the reservoirs' dead storage volume. It was regularly observed, though, that active storage siltation initiated far earlier and fiercer than scheduled, i.e. before the complete dead storage volume was exhausted (Reza Rahmanian & Banihashemi, 2012; Basson, 1999). Consequently, this research represents a non-realistic optimal state from the reservoir management perspective. The actual reduction of active storage must be considered to be fiercer than suggested in this study (Kantoush & Sumi, 2017).
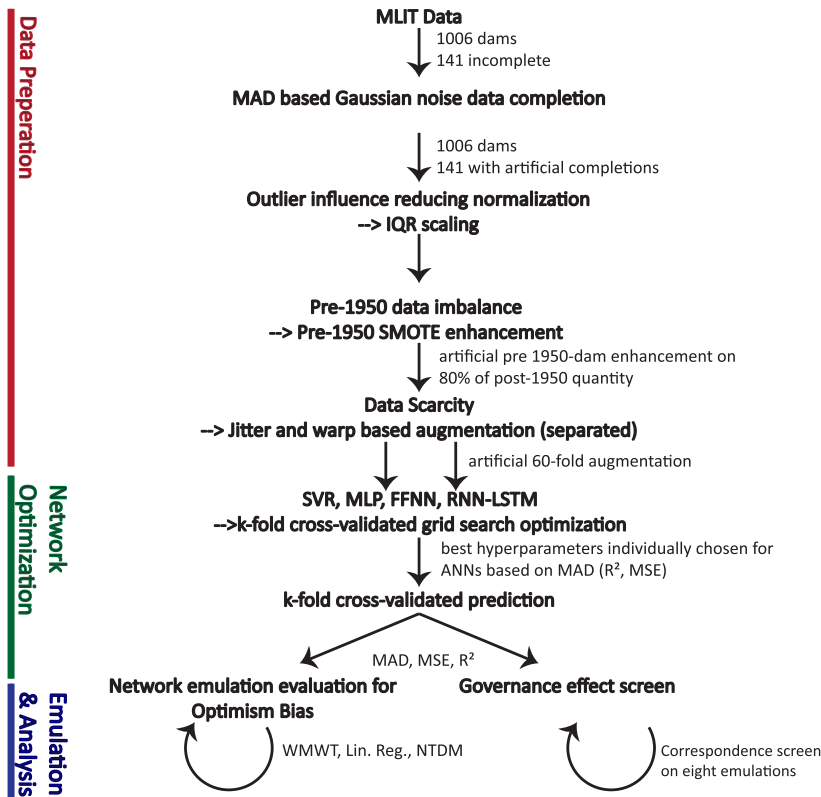
MLIT Data
1006 dams
141 incomplete

MAD based Gaussian noise data completion

1006 dams
141 with artificial completions

Outlier influence reducing normalization
--> IQR scaling

Pre-1950 data imbalance
--> Pre-1950 SMOTE enhancement
artificial pre 1950-dam enhancement on
80% of post-1950 quantity

Data Scarcity
--> Jitter and warp based augmentation (separated)
artificial 60-fold augmentation

SVR, MLP, FFNN, RNN-LSTM
-->k-fold cross-validated grid search optimization
best hyperparameters individually chosen for
ANNs based on MAD (R^2, MSE)

k-fold cross-validated prediction

MAD, MSE, R^2

Network emulation evaluation for
Optimism Bias
WMWT, Lin. Reg., NTDM

Governance effect screen
Correspondence screen
on eight emulations

Data Preperation

Network Optimization

Emulation & Analysis

Figure 2. Flowchart of the methodological process.

If arranged by the inauguration year, the base data represent a time series that allows us to draw conclusions about general siltation management performance in Japan. The terms *general* or *average* refer in this context to siltation performance behaviour that is to be derived from the Japan overarching conglomerated data by the ANNs. The emulations shall thus produce a siltation development archetype that is typical for Japan based on the available data (approach).

To do so, the data were transformed for ANN purposes and streamlined to the desired key information: 1. The year of reservoir inauguration and 2. the reported lake siltation versus the dead storage volume.

2.3. ANNs

For the given data, a non-linear relation is presumed. ANNs are considered a good tool for the analysis, pattern detection and data emulation for this type of data behavior (Gamboa, 2017; Michelucci, 2018; Schmidhuber, 2014). First, simple network attempts were already conceived and approached throughout the 1960s, 1970s and 1980s. A breakthrough of the ANN methodology was reached in the recent decade due to a significant increase in CPU and GPU computation power (Chellapilla, Puri, & Simard, 2006), which allowed ANNs to grow in size and solve complex problems. Due to the amount of computed neurons

and the necessary CPU/GPU power, the ANN method is described as a brute force approach (Perez, 2017) or reinforcement learning (Zhang, Vinyals, Munos, & Bengio, 2018).

A well-trained ANN is capable of emulating, forecasting or even creating realistic results based on limited or incomplete non-linear data (Michelucci, 2018; Schmidhuber, 2014). This situation is given for the Japanese (or every other country's) siltation case.

ANNs are thus a viable tool to emulate a generalized contemporary state of siltation for reservoirs of certain ages in Japan. The ANN mass data analysis and will thus give feedback and insights on the Japan wide overall state of reservoir siltation and the presumed deviations from the actually planned siltation state. This allows to draw conclusions on large scale large infrastructure management of a country and hence the (not-)perpetuation of sustainability according to the World Commission on Environment and Development (2018) definition.

ANNs consist of different layers of weighting entities, so-called neurons. Their objective is to emulate a variety of known relations that are presented to them based on training data. The neuron entities iteratively receive and process the data within different kinds of internal connections whilst perpetually updating their weights. Performance is checked on a separate validation test set. The neurons are designed to reduce the asserted deviation between the known training data relation and the results of the own emulation of the prior iteration by a certain, determined degree (Michelucci, 2018; Schmidhuber, 2014).

As the relation of training and validation data is known, the emulation accuracy on test and validation data is used as a guidance on network performance. In a concluding network performance check, the ANNs capability to generalize on new data is surveyed by an iteration on the hitherto unseen test data (Michelucci, 2018; Schmidhuber, 2014).

To suit ANN demands, the data need to be randomly split into a training part and a testing part. In this study, this split is done by k-fold cross-validation. The technique separates the complete data into k entities, out of which k minus one form the training part. The left k-sized fraction forms the test set used for network validation. In the next iteration, the network is trained and tested on another k-set. This methodology equals out extremes caused by arbitrary chosen splits. Moreover, it guarantees the usage of the full test set. It is also used for optimization purposes (s. next chapter) (Jiang & Chen, 2016).

ANNs face several issues that reduce their performance or contort results. Among them are overfitting (Allamy, 2014; Zhang et al., 2018) and underfitting (Allamy, 2014), data scarcity, the need for normalization, data imbalance and outlier influence (Khamis, Ismail, Khalid, & Tarmizi Mohammed, 2005). These issues were addressed using methods such as dropout (Park & Kwak, 2017), augmentation (jitter (pure Gaussian noise) and warp (Gaussian noise on Bezier-Curves))(Le Guennec, Malinowski, & Tavenard, 2016; Um et al., 2017; Velasco, Garnica, Lanchares, Botella, & Ignacio Hidalgo, 2018; Xiao & Xu, 2012), synthetic minority oversampling technique (SMOTE) (Fernández, García, Herrera, & Chawla, 2018), interquartile range (IQR) scaling (Mizera et al., 2004) and median absolute deviation (MAD) (Gorard, 2013) based Gaussian noise data completion. The complete process is shown in Figure 2. The jitter and warp-based augmentations were carried out separately and displayed discretely in the results. Warp is used to produce better results, as the enhancement follows a more continuous process.

2.4. Selection of network hyperparameters

Several ANN types are applicable for time series emulation, including support vector machines for regressions (SVR) (Nisbet, Miner, & Yale, 2017), feed forward networks (FFNN) or multilayer perceptrons(MLP) (Michelucci, 2018), and long short term memory recurrent ANNs (RNN-LSTM) (Michelucci, 2018). MLPs and FFNNs resemble each other quite lot. Thus, their emulations are expected to be similar.

SVRs and MLPs were applied via SKlearn libraries (Pedregosa et al., 2011), FFNNs and RNN-LSTMs were constructed using TensorFlow and Keras libraries (Abadi et al., 2015).

ANNs succumb to hyperparameters that regulate them. These hyperparameters have tremendous influence on their performance (Michelucci, 2018; Nisbet et al., 2017). To select the optimal operation environment, the best-performing hyperparameters were chosen using an iterative k-fold cross-validation grid search process (Jiang & Chen, 2016). The optimization guidance was formed using R^2, MAD and mean squared error (MSE) as metrics. Thus, MAD was the prime criterion because it is comparatively robust to outlier distortion in a non-linear environment (Gorard, 2013). The other metrics were used if they improved by 15% after a cross-validation iteration. The optimized values were chosen as hyperparameters, which can be seen in Table 1. The emulations themselves were also carried out in a k-fold cross-validated manner. Overall, eight emulations were produced.

2.5. Network metrics

Over- and underfitting performance were originally considered as a measure to select the best-performing two ANN types, since over- and underfitted ANNs are not capable to generalize appropriately on new data. Such networks either emulate the testing data in an overly exact ragged fashion (overfitting) or fail to react to each type of new data (underfitting) (Allamy, 2014; Zhang et al., 2018). The selection process was planned to be carried out via the analysis of R^2 performance. Due to reasons stated in the chapter "Metrics' Validity" in the Discussion, no best-performing network was selected and all four are displayed as main results.

Their evaluation is carried out via MSE, R^2 and MAD. The focus lies on the latter, as it is more robust against the influence of outliers (Gorard, 2013), which is important

Table 1. Overview of k-fold cross-validated grid-searched hyperparameters. Please refer to the text for more details on abbreviations, etc.

Hyperparameter	Network Specific Optimized Setups			
	SVR	MLP	FFNN	RNN-LSTM
Learning Rate	-	0.0006	0.015	0.004
Batch Size	-	128	100	256
Optimizer	-	Adam	Nadam	Adam
Loss Function	-	-	mae	logcosh
Activation	-	tanh	relu	tanh
Epochs	-	automatic	15	12
Dropout	-	-	10% per layer	50% per layer
Hidden layer setup	-	500/150/70/25/1	500/150/200/50/1	250/120/10/1
k-fold emulation	4	4	8	3
C	1.20E+02	-	-	-
Gamma	1	-	-	-
Epsilon	0.004			
Kernel	rbf	-	-	-

considering the generalization approach of the study. In nonlinear relations, however, the metrics have a limited significance. This is discussed later in "Metrics' Validity".

Furthermore, a mere linear regression was computed from the original data to test whether major deviations between the elaborated ANN approach and the linear regression exists. The linear regression thus serves as an indication of whether the ANN approach is actually worthwhile.

2.6. Identification of an optimism bias

The networks were evaluated regarding optimism bias in a way similar to prior research by Ansar et al. (2014). The Wilcoxon-Mann-Whitney Test (WMWT) was used to identify whether the optimism bias hypothesis should be rejected (Corder and Foreman, 2014).

2.7. Identification of governance impact

One of the aims of this research is to substantiate a presumed sustainable effect of governance on sediment management.

The research focuses on two elements: 1. legal and administrative directives and 2. the economic environment. For the prior, regulatory means for indication and monitoring were chosen as prevalent. For the latter, the procurement competition situation was selected.

A substantiation is assumed if an ANN-identified period of reduced siltation accumulation coincides with periods of governance intervention in Japan. Those intervention periods are to be identified.

2.7.1. Indication and monitoring

The implementation of the *Kasen Hobou Gijutsu Kihon* – 河川破防技術基準(Norm for relevant techniques against destructive river basin management; NTDM) (Okano et al., 2004) in 1957 (Shouwa 32 – 昭和32) was a major incident that changed the approach of indication and monitoring for reservoirs within Japan. Another contribution of that epoch is a revision of the NTDM. It enforces systematic monitoring and legal pressure and is reported for the middle of the 1960s (Okano et al., 2004). The years around 1957 form a period of special interest.

2.7.2. Competition

Following Flyvbjerg and Ansar, a non-competition situation in public procurement leads to a higher observed optimism bias and thus to irregularities (such as inappropriately managed siltation) in negative effects for mega-projects. Competition should be seen as a contribution towards an aim-based project management (Ansar et al., 2014).

In Japan, several events were highly influenced by competition environment in public procurement economics (Fujii, 2012). Prior to the second World War the *Shimei Kyousou* (指名競争) was prevalent. Business collusion between the state as tenderee and its contractor, the powerful Japanese *Zaibatsu* (財閥), was a daily occurrence. Competition-induced pressure for performance increases was de facto non-existent (Fujii, 2012).

In 1947 competition was pushed by the *Shiteki Dokusen no Kinshi oyobi Kousei Torihiki Kakuho ni kan suru Houritsu (Dokusen Kinshi Hou)* – 私的独占の禁止及び公正取引確保に

関する法律(独占禁止法), the so-called Antimonopoly Act. Due to a lack of control and political unwillingness, however, the effect was limited (Kinoshita, Satou, Matsumoto, Tanaka, & Tanno, 2008). The rather non-competitive state continued to be in effect for public construction tendering, as the 1981-Shizuoka incident demonstrated. It evolved around the ministry of construction and its then minister Saitou Shigeyoshi, who had to step back after large scale bid-rigging was revealed to the Japanese public (Fujii, 2012; Kinoshita et al., 2008). Kinoshita (2012) confirms that during this era the minority of tenders were subject to real competition.

The situation changed in 1994 with the *Koukyou Jigyou no Nyuusatsuãƒ»Keiyaku Tedzuzuki no Kaiyen ni kan suru Koudou Keikaku* – 公共事業の入札・契約手続の改善に関する行動計画 (Public procurement: Plan for the improvement of contraction procedures)., which was the most important step in the public procurement economy since 1900 (Kinoshita et al., 2008). The years around 1947 and 1994 are thus important time period for improved competition. Among both events, 1994 had a higher influence.

3. Results

The produced emulations reveal a uniform picture. The siltation is fiercer than the design demanded by law (Okano et al., 2004). The median years for a certain fraction of siltation can be derived from Table 2. The divergence between some of the emulations is a matter for the discussion part.

For 100% dead storage siltation the median of the eight simulations lays in the year 1964, i.e. reservoirs built in that year are generally expected to have lost their complete dead storage volume. For 50% siltation, the median emulation expects the year 1985, i.e., dams constructed in that year are assumed to have lost half of their dead storage volume. All emulations reach 150% siltation of dead storage, some far more.

Three ANN types show similar behavior with several distinguishable phases, which will be described below. LSTM emulations present an excerpt with a rather continuous siltation growth emulation.

In contrast, the scale of three ANN types is rather uniform, with MLP types being the exception. MLP types project a considerable fierce emulation of siltation growth. The entire range of emulations can be found in Figure 3. The variability that the generalized ANNs deem likely is denoted by ragged parts and the boldness of strokes.

The metrics for most emulations show uniform patterns in Table 3. MLP emulation metrics diverge from the rest, which corresponds to their different scale. LSTM metrics also differ slightly from FFNNs and SVRs. The discussion will address the implications of these metrics.

Table 2. Median siltation of dead storage emulations.

| Dead Storage Siltation	Jitter				Warp				Median
	SVR	FFNN	LSTM	MLP	SVR	FFNN	LSTM	MLP	
50%	37	33	17	26	38	38	24	24	29.5
100%	54	52	48	41	53	51	51	38	51
150%	101	81	88	49	81	80	87	53	81
200%	-	-	-	72	104	109	-	70	-
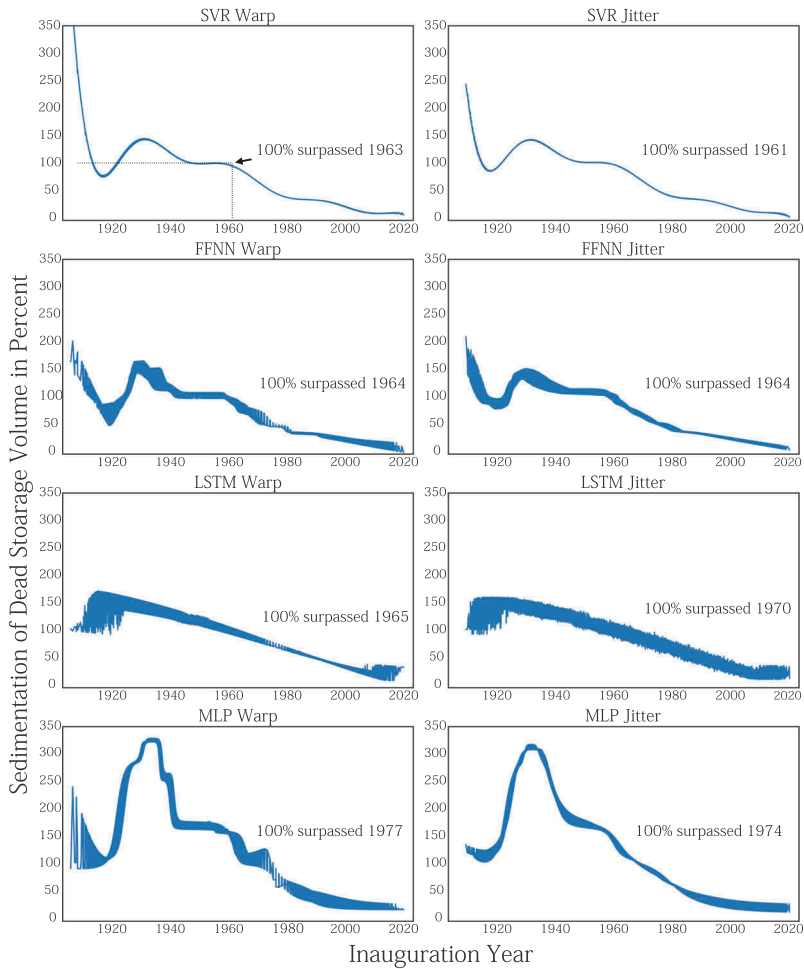
Figure 3. Eight emulations based on four ANN types and two augmentation methodologies.

Table 3. Metrics of the eight emulations.

Metric	Jitter				Warp			
	SVR	FFNN	LSTM	MLP	SVR	FFNN	LSTM	MLP
MAD	0.3912	0.3754	0.4136	0.6521	0.3793	0.3654	0.4023	0.6385
MSE	9.99	10.07	9.893	9.154	10.14	10.22	10.05	9.238
R^2	0.0127	0.0074	0.0246	0.0965	0.0139	0.0099	0.0259	0.1011

From Figure 4 it is evident that most emulations differ significantly from a mere linear regression. In fact, the linear regression indicates a far fiercer siltation rate for the reservoirs than the ANN emulations. Only the MLP emulations range in the order of the same increment.

The concordance (in shape and scale) of the emulations insinuates that a mere linear regression cannot reflect the complex non-linear behavior of a long-time reservoir siltation, i.e., that the ANN methodology is worthwhile.
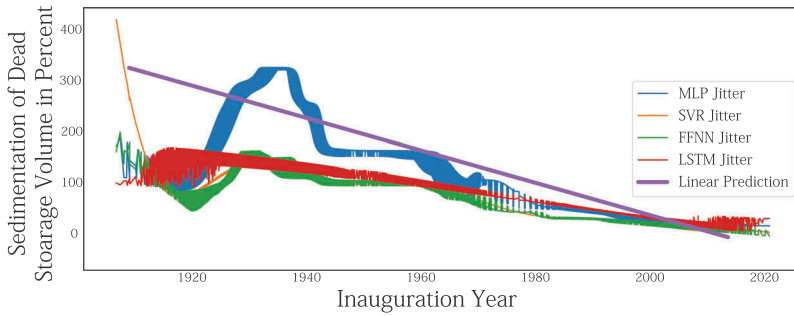
Figure 4. Warp augmented emulation in comparison to a linear regression.

The subtraction of the NTDM from the network emulations reveals oversiltation behavior, which is demonstrated by the equation below (OS = Oversiltation, PS = Projected Dead Storage Siltation in %, DS = Dead Storage Siltation demanded by the 1957 Directive in %).

$$OS = PS - DS_{2015} \qquad (1)$$

Oversiltation is prevalent for all types of emulations, as shown by Figure 6. Based on the shape characteristics, it can be sub-classified into five different phases whose names are based roughly on historic or Japanese epochs. An exemplary breakdown can be consulted in Figure 5. The enumeration is carried out from present to past from right to left on the graph.

(1) *Heisei Antei* – 平成安定.Heisei Stability: Constant oversiltation, ranging around 20% up to year 1980
(2) *Shouwa Kakou* – 昭和下降.Shouwa Drop: A drop of oversiltation towards the Heisei-level, initiated from the end of the 1950s/beginning of the 1960s.
(3) *Shouwa Antei* – 昭和安定.Shouwa Stability: Rather constant level of over-siltation from the mid-1940s onward that ranges approximately 60% for FFNN and SVR (and LSTM) and 100% for MLP.
(4) *Sekai Sensou Kakou* – 世界戦争下降. A drop of a very steep (MLP) or rather smooth nature during the time of the second World War.
(5) *Taishou Aimai* – 大正曖昧. Taishou Ambiguity: Values of high ambiguity, roughly around the Taishou time.
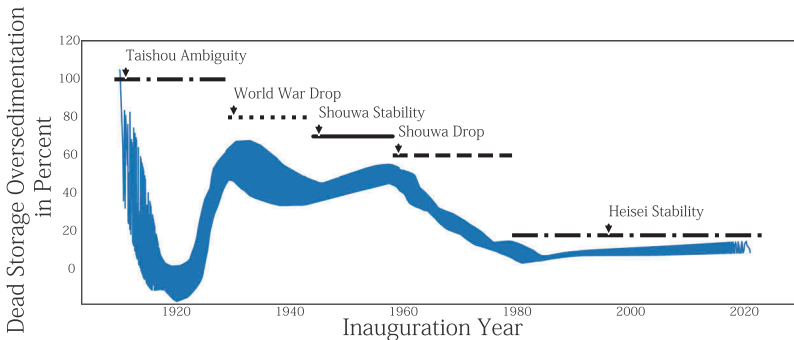
Figure 5. The Jitter-augmented FFNN emulation and the identified periods of siltation behavior.
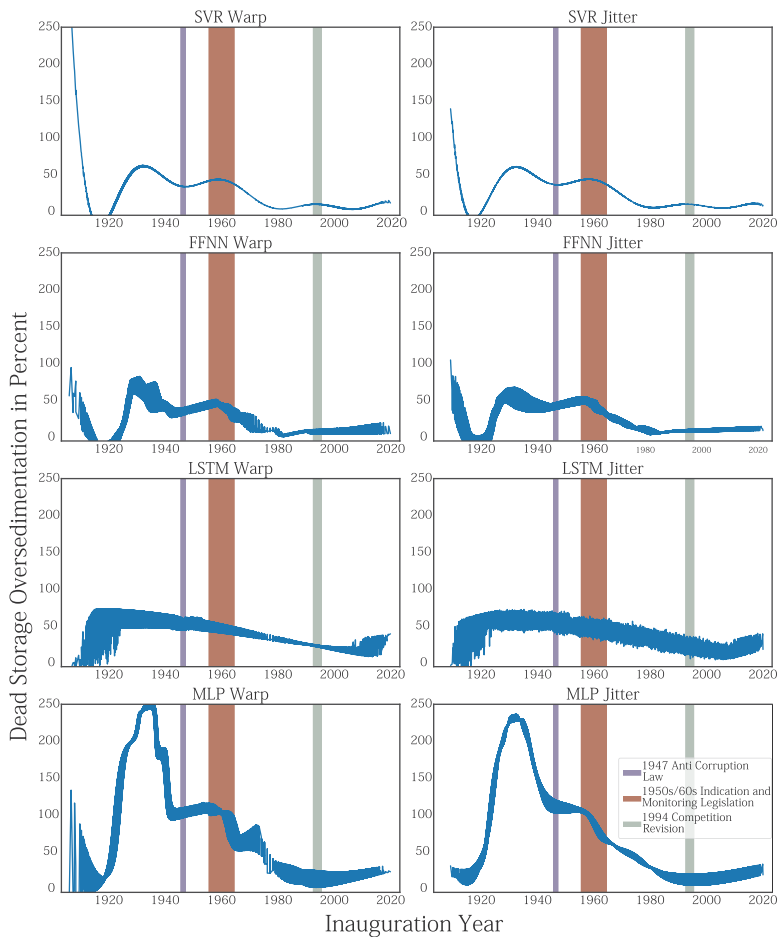
Figure 6. Degree of oversiltation in all emulations and coincidence with governance interventions with a potential impact on siltation management performance.

Two periods reflect a time of significant oversiltation reduction, i.e. an im- provement of siltation treatment and management. One coincides for all emulations of a characteristic shape with one of the three periods of identified governance intervention: The NTDM laws were established when the Shouwa Drop was triggered. as seen in Figure 6. It is an indication of the contribution of governance to improved siltation management.

Because siltation of the general Japanese dam lies constantly beyond the 1957-demanded NTDM range, the presumption is that planers and managers were not capable of meeting its standards, including in recent projects, as Figure 6 demonstrates.

A comparison in Figure 7 with the siltation values of six coincidentally elected reservoirs that were recorded between 2000 and 2017 shows furthermore that a rough quite a concordance in shape and behaviour with the Jitter augmented FFNN and thus with the five identified phases. It is emphasized that the data of the six reservoirs (provided by the PWRI and MLIT) were not included in the ANN base data, i.e. that it is unknown to the ANN. This shows that the ANN have achieved quite a good generalization on Japanese
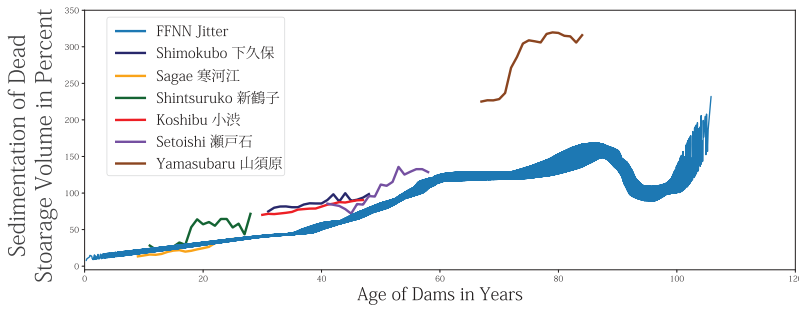
Figure 7. Comparison of observed and coincidentally selected dead storage siltation values with emulated dead storage values.

dam behaviour given that the selected dams originate from different regions, are different size and inaugurated in different decades.

It is furthermore assumed that the reservoir projects would not have been permitted if authorities, planners and managers were not convinced to meet the NTDM. Therefore, systematic conduct of high optimism can be derived.

The WMWT supports this by assigning every ANN emulation statistical significance with p-values close to zero and U-values far less than the critical U (e.g. Warp-augmented FFNN: p = 2.39e-182 and U = 32774307 far less than the U_crit= 57623970.173) when compared to a distribution of values that meets the NTDM. The null hypothesis is consequently rejected.

This means, the emulated values and values that meet the NTDM are not from the same distribution and they indeed represent the outside view demanded by Kahneman and Tversky (1979); Flyvbjerg (2016). As already demonstrated in Figure 6, the values are way higher than NTDM values. Hence, emulated values derived from real data are not expected to meet the NTDM. This happens – as the WMWT shows – in a statistically significant and systematic manner. With regard to fullfilling the NTDM, it must be assumed that generally both siltation was underestimated and the own skills overestimated. This implies that siltation management suffers broadly from optimism bias.

4. Discussion

The ANN emulations differ in shape and range, which should be expected, as they are influenced by different build-ups (Michelucci, 2018; Nisbet et al., 2017). Interestingly, many similarities exist among the emulations. At some points, differences were noted where they were not expected.

4.1. The divergence of MLPs and LSTMs

The difference in the scale of the MLP emulation towards the other network emulations is striking. The shape, however, resembles that of SVRs and FFNNs. Originally, FFNNs and MLPs were rather similar, as their internal structures are based on a very similar feedforward approach (Makrem, Imen, & Kaïs, 2016). The different results can most likely be

explained by two features: the Dropout (Park & Kwak, 2017) and the Loss (Li, Xu, Taylor, Studer, & Goldstein, 2018), that were missing in the applied SKlearn architecture for MLPs.

The LSTM shape divergence has to be attributed to the completely different build-up of LSTM-Neurons, which are designed to have a memory function (Michelucci, 2018). LSTMs are thus more robust to short-run deviations compared to ANNs (Shah, Campbell, & Zulkernine, 2018). However, it is not entirely clear why this phenomenon causes the observable effect.

4.2. Metrics' validity

As seen in Table 3, MLPs appear to outperform the other ANNs in terms of the MSE and the R^2 whilst being themselves on a rather mediocre level in absolute terms. However, the significance of such metrics needs to be discussed.

First, it is important to emphasize that the comparability between them is partly limited because the emulations succumb to differently randomized k-fold splits of the same base data.

Second, according to Spiess and Neumeyer (2010); Alexander, Tropha, and Winkler (2015); Pontius, Thontteh, and Chen (2008) the validity of both R^2 and RMSE is questionable in non-linear environments. Achen (1990) argues that their whole significance is rather limited. Nevertheless, they are used quite regularly (McKellar & Lu, 2003).

In this study, a non-linear ANN emulation, especially R^2 is viewed under a completely different angle, namely as an indication of overfitting. Under non-linear deep learning circumstances, an optimum value for both R^2 and RSME is not desirable, which would mean an absolute fit and thus non-generalization, i.e., overfit. In contrast, low performance indicates a complete no-fit and thus an underfit. Acceptable but imperfect R^2 and RMSE results are hence desirable. Because this approach is not entirely verified, no ANN type was omitted during this study due to supposed overfitting likeliness.

Overall, the MAD is seen as a more valid indication of a good emulation fit. Although the MAD basically resembles the MSE, the focus on the median reduces outlier influence. Emulations that do not cover all outliers receive a better score, as reflected in Table 3; they appear to be more generalized (Gorard, 2013) and thus of higher scientific value.

4.3. Impact of augmentation technique

Although the emulation results differ, a clear recommendation on which tech nique produces better results cannot be drawn. Other than expected, Warp-augmented emulations are neither clearly distinguishable nor bear significant advantages to Jitter-augmented emulations. The demonstrated closeness of re-sults, however, points towards the robustness of the ANN approach.

4.4. The Taishou ambiguity and the world war drop

Data imbalance has a negative impact on ANN performance (Michelucci, 2018) which exists for pre-WWII data. Almost no reservoirs were constructed during that period of time as Figure 8 demonstrates. Their data footprint is thus amplified compared to that for dams in other epochs.
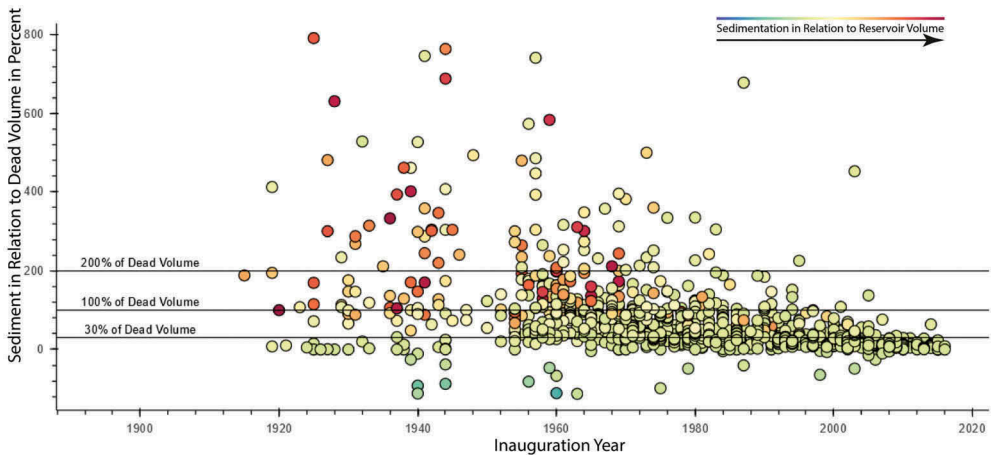
Figure 8. Distribution of siltation in Japanese reservoirs based on 2015 data.

Several measures were undertaken to reduce the data imbalance of pre-WWII data. Though improvements were made, the Taishou Ambiguity demonstrates that these imbalances were not overcome entirely. A rapid decline and a subsequent rise in dead storage siltation is observable for many emulations.

This finding should not be taken literally; rather, it is a directional cue. If the drop around 1920 were omitted, most emulations return to the prior magnitude of oversiltation, as seen in Figure 6. The emulation graphs would then comply with the asymptotic saturation behavior of siltation found in other research (Annandale, 2013; Basson, 1999; Kantoush & Sumi, 2017).

## 4.5. Japanese reservoir situation and intrinsic optimism bias

According to Table 2 the dead storage capacity for a contemporary dam is likely to be consumed after 51 years of operation. Figure 8 demonstrates the intensity of the Japanese reservoir situation.

With respect to the inauguration distribution seen in Figure 9, this finding is quite alarming. The majority of large and very large dams will soon reach 100% capacity of dead storage volume, whilst the medium and small dams will likely do so within the next 5 to 10 years.

As mentioned in the introduction, it is needless to say that this analysis does not account for every single dam – there are huge individual differences. Nevertheless, if the Japanese reservoirs were regarded as an entity, the result is a reason for profound concern regarding reservoir sustainability.

Because Japan is among the best-performing countries in terms of siltation management(Auel et al., 2016; Kondolf et al., 2014), this finding implies an even more somber situation for other countries.

The fact that the dead storage volume is comparatively small in relation to the rest of the reservoir, as shown in Figure 10, means no relief for two reasons:
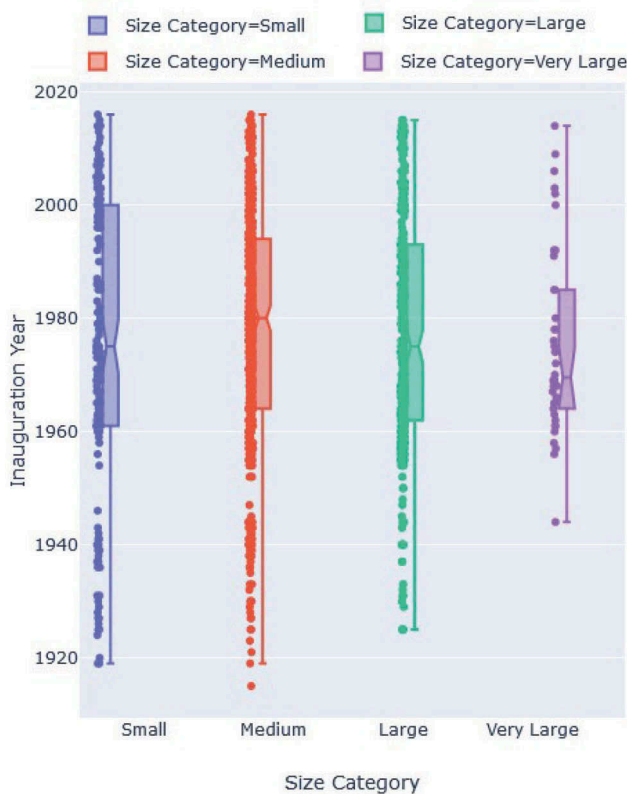
Figure 9. Distribution of dam inaugurations and corresponding medians; small = 1975, medium = 1980, large = 1975, very large = 1969.

(1) As (Annandale, 2013; Basson, 2009; Kantoush & Sumi, 2017) imply, the siltation of a reservoir does not occur entirely within the dead storage but also affects a considerable fraction of active storage. Because this study found that in most reservoirs, siltation is greater than expected, and the fraction of reduced active storage is also assumed to be greater than expected.

(2) The dead storage serves as a marker. If clear laws such as the NTDM cannot be kept, the implication can be made that longevity is a feature that will not be kept. Thus, it can be inferred that most reservoirs are not constructed in sustainable way, but rather in an optimistic way due to the inside view (Kahneman, 2011). They are optimistically biased. As Japan is a pioneer in sediment management (Kantoush & Sumi, 2019; Kondolf et al., 2014; Sumi, 2015), we can assume that the effect is stronger throughout the rest of the world, which implies that the global sustainability of reservoirs is not guaranteed.

4.6. The impact of governmental action

The introduction of the Antimonopoly Act and, hence, the effect of induced competition on siltation in 1947 is not clearly perceivable. Moreover, almost no effect is noticeable for the enforcement of competition in 1994. Among the 1947 and 1994 acts, the latter has
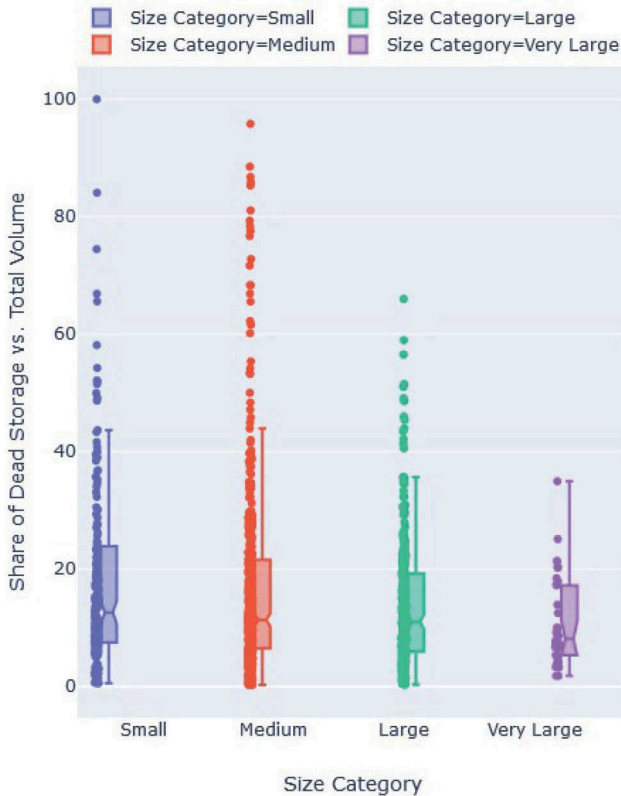
Figure 10. Fraction of dead volume in relation to total reservoir volume with corresponding medians. Small = 12.63%, medium = 11.35%, large = 11.03%, very large = 8.13%.

a higher impact on competition according to Kinoshita et al. (2008) Thus, it seems likely that competition has had no impact on siltation performance.

Two assumptions regarding this lack of effect might serve as explanations. First, siltation is of minor significance for the economic competition that companies the planning and construction of reservoirs, which fits the globally observed pattern (Annandale, 2013). If long-term siltation performance is a decisive criterion in reservoir economics, far more new dams would include new measurements for reduced siltation (Kantoush & Sumi, 2017; Zarfl et al., 2015). Second, the laws were of minor efficiency, as indeed reported by Kanda and Fujii (2017): Dumping in pricing and business collusion are still quite established in Japan. Thus, competition might not have any effect because it is still nonexistent.

Contrarily to the competition, the introduction of strict obligations regarding monitoring and indication seem to have influenced sediment performance. As Figure 6 shows, the oversiltation reduction during the Shouwa Drop coincides for all of the emulations quite neatly with NTDM introductions after 1957. Following the upper observations, the perceived governmental effects must be assigned to the obligations made for siltation monitoring and indicating.

5. Further thoughts on the variety of siltation conditions

The hydrological variability caused by climate changes have an increasingly high impact on water and sediment inflow to the reservoir. One of the main shortcomings in global siltation management is the implementation of a widespread quantification of flow and sediment rating curves under the changing climate.

Annandale (2013) explained that the magnitude and spread of the anticipated increase in hydrologic variability is not known. Therefore, it is necessary to make both engineering and governmental decisions to deal with the uncertainties.

Kantoush and Sumi (2019) explained that there are various challenges regarding dam management due to climate changes and increased number of flood peaks. A periodical monitoring for reservoir siltation before and after every flood event is seen as a very valuable practice. In case of increased reservoir siltation, an immediate intervention or an upgrading of the dam facilities is necessary to recover the original dam functionality and to ensure reservoir capacity is recommended to cope with the fiercer flooding events. Moreover, physical processes of sediment transport and siltation as well as upgrading and retrofitting effects on reservoirs of high age need to be understood to guarantee a sustainable long-term operation,

Measuring the benefits and costs of siltation management improvement for the development of a sustainable infrastructure is often difficult. However, it is unavoidable that siltation management needs to perform economic feasibility studies for the methodologies in consideration.

The long-term economic, ecologic and social costs and benefits of asset management in coordination with changes in dam and downstream should be in perspective during decision making (Annandale, 2013; Kantoush & Sumi, 2019). This should extend to include a thorough assessments of climate change impacts on aging reservoirs and the determination of ecosystem responses to new reservoir projects as well as to the ongoing loss of reservoir functionality of already installed dams. Moreover, critical studies to demonstrate the social dimension of man-made reservoir interventions is inevitable for adequate sediment management.

From the authors point of view and based on the findings regarding the optimism biased management, (new) concepts and methodologies should be conceived a priori to design intergenerational, sustainable, self-supporting rehabilitation systems for river basins with man-made reservoirs. Further research is needed to guide the future management for aging Japanese and global reservoirs to optimize and reflect both the huge investment decisions that will have to be made and the legacy that reservoirs mean for the future generations.

6. Conclusions and outlook

The applied ANN methodologies serve as a powerful tool for detecting optimism bias. The latter likely exists in Japanese reservoirs because compliance with the NTDM directive could not be assured. This observation implies that planners and practitioners hazard the consequences consciously or unconsciously.

Although reservoir designers have powerful techniques at hand, they should consider a more conservative approach for dam longevity and actively seek for outside view information like the ones of the present paper.

Neither the omission nor the implication of *competition* as a tool led to a different sediment behavior during Japanese history. It is thus presumed that the reforms had limited effects and/or that siltation does not play a decisive role in the consideration of public procurement. This behavior is assumed to be global. In this case, it would have a devastating impact on the worldwide long-term sustainability of reservoirs.

Emphasis is given to the power that governmental approaches possess, if applied correctly (with an incorporated outside view). Countries and regions that are confronted with sediment issues should take this power into consideration. Dams constructed under such new directives might not produce the highest revenue in the short run but will most certainly do so in the long run.

Moreover, active management and refurbishment methodologies should be undertaken to guarantee and ensure the longevity and sustainability in already constructed and new dams, as Kantoush and Sumi (2017); Auel et al. (2016) demonstrate.

Upcoming in-depth multivariate, multi-time series approaches with ANN methodologies will produce more detailed findings which should be useful in gaining more insight for siltation management in Japan and worldwide.

Acknowledgments

Declaration of originality

I declare that this research is my own work except where there is clear acknowledgement 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

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., … Zheng, X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software Retrieved from tensorflow.org.

Achen, C. H. (1990). What does "explained variance"explain?: Reply. *Political Analysis*, *2*, 173–184.

Alexander, D. L. J., Tropha, A., & Winkler, D. A. (2015). Beware of R2: Simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. *Journal of Chemical Information and Modeling*, *55*(7), 1316–1322. doi:10.1021/acs.jcim.5b00206.

Allamy, H. (2014). Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study). *Computer Science, Communication & Instrumentation Devices*. doi:10.3850/978-981-09-5247-1_017

Annandale, G. (2013). *Quenching the thirst: Sustainable water supply and climate change*. Scotts Valley, CA: CreateSpace Independent Publishing Platform.

Ansar, A., Flyvbjerg, B., Budzier, A., & Lunn, D. (2014). Should we build more large dams? The actual costs of hydropower megaproject development. *Energy Policy*, *69*(C), 43–56. Retrieved from https://EconPapers.repec.org/RePEc:eee:enepol:v:69:y:2014:i:c:p:43-56

Auel, C., Kantoush, S., & Sumi, T. (2016). *Positive effects of reservoir sedi- mentation management on reservoir life - examples from Japan*. 84th ICOLD Annual Meeting, At Johannesburg, South Africa.

Basson, G. R. (1999). *Dealing with reservoir sedimentation: Guidelines and case studies*. Paris: Bulletin (Commission Internationale des Grands Barrages) Commission International des Grands Barrages.

Basson, G. R. (2009). *Management of siltation in existing and new reservoirs*. 23rd Congress of the International Commission on Large Dams CIGBICOLD, Brasilia.

Chellapilla, K., Puri, S., & Simard, P. (2006). *High performance convolutional neural networks for document processing*. La Baule, France: Suvisoft.

Corder, G. W., & Foreman, D. I. (2014). *Nonparametric statistics. A step-by- step approach* (2nd ed.). Hoboken, NJ: Wiley.

de Vente, J., Poesen, J., Verstraeten, G., Govers, G., Vanmaercke, M., Rompaey, A. V., … Boix-Fayos, C. (2013). Predicting soil erosion and sediment yield at regional scales: Where do we stand?. *Earth-Science Reviews, 127*, 16–29.

Fernández, A., García, S., Herrera, F., & Chawla, N. V. (2018). Smote for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary. *Journal of Artificial Intelligence Research*, *61*(1), 863–905.

Flyvbjerg, B. (2006). From Nobel Prize to project management: Gettings risks right. *Project Management*, *37*, 5–15.

Flyvbjerg, B. (2016, April). How to get your business case right, by using Kahneman's outside view. Retrieved from http://dx.doi.org/10.2139/ssrn.2769829

Fujii, A. (2012). *Japan will perish due to Compliance: The Assets and Drawbacks of the New Liberalism* [コンプライアンスが日本を潰す: 新自由主義との攻防]. Tokyo: New Releases of the Fusou Publishing House [扶桑社新書]. Fusousha [扶桑社].

Gamboa, J. C. B. (2017). *Deep learning for time-series analysis*. Retrieved from https://arxiv.org/abs/1701.01887

Ghimire, G., & DeVantier, B. (2016). Sediment modeling to develop a deposition prediction model at the olmsted locks and dam area. *World Environmental and Water Resources Congress 2016*, 410–420.

Gorard, S. (2013). The possible advantages of the mean absolute deviation 'effect'size. *Social Research Update*, *65*, 1.

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.

Jiang, P., & Chen, J. (2016). Displacement prediction of landslide based on generalized regression neural networks with k-fold cross-validation. *Neuro- Computing*, *198*, 40–47. Advances in Neural Networks, Intelligent Control and Information Processing.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus and Giroux.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *2*, 263–292.

Kanda, Y., & Fujii, A. (2017). Study on the bid tendering and contracting system for public works considering public benefit. *Policy and Practice Studies* [実践政策学], *3*(1), 39–52.

Kantoush, S., & Sumi, T. (2010). River morphology and sediment management strategies for sustainable reservoir in Japan and European alps. *Annuals of Disaster Prevention Research Institute*, *53*, 82–89.

Kantoush, S. A., & Sumi, T. (2017). *The aging of Japan's dams: Innovative technologies for improving dams water and sediment management, river sedimentation*. London: Taylor and Francis.

Kantoush, S. A., & Sumi, T. (2019). Paradigm shift for sediment man- agement. Retrieved from https://www.waterpowermagazine.com/features/featureparadigm-shift-for-sediment-7263873/

Khamis, A., Ismail, Z., Khalid, H., & Tarmizi Mohammed, A. (2005). The effects of outliers data on neural network performance. *Journal of Applied Sciences*, *5*, 1394–1398.

Kinoshita, M. (2012). *Public Procurement Research: Why the creation of a reasonable competition environment is something the world has not yet seen* [公共調達研究: 健全な競争環境の創造に向けてなぜ, 世界に例をみない制度になったか]. Tokyo: Daily Construction Works News Journal [日刊建設工業新聞社].

Kinoshita, M., Satou, N., Matsumoto, N., Tanaka, Y., & Tanno, H. (2008). Investigation regarding the transitions and contemporary conditions of the tendering and contraction system in public construction [公共工事の入札契約制度の変遷と今後のあり方に関する考察]. *Collected Research of Construction Management* [建設マネジメント研究論文集], *15*, 289–300.

Kondolf, G. M., Gao, Y., Annandale, G. W., Morris, G. L., Jiang, E., Zhang, J., ... Wei, Z. (2014). Sustainable sediment management in reservoirs and regulated rivers: Experiences from five continents. *Earth's Future*, *2*, 256–280.

Le Guennec, A., Malinowski, S., & Tavenard, R. (2016). Data augmentation for time series classification using convolutional neural networks. In *ECML/PKDD workshop on advanced analytics and learning on temporal data*. Italy: Riva Del Garda.

Li, H., Xu, Z., Taylor, G., Studer, C., & Goldstein, T. (2018). Visualizing the loss landscape of neural nets. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, & R. Garnett (Eds.), *Advances in neural information processing systems 31* (pp. 6389–6399). New York, NY: Curran Associates, Inc.

Makrem, B. J., Imen, J., & Kaïs, O. (2016). Study of speaker recognition system based on feed forward deep neural networks exploring text-dependent mode. In *2016 7th International Conference on Sciences of Electronics, Tech- nologies of Information and Telecommunications (SETIT)* (pp. 355–360). Hammamet, Tunisia.

McKellar, R., & Lu, X. (2003). *Modeling microbial responses in food*. Contemporary Food Boca Raton, FL: Science CRC Press.

Michelucci, U. (2018). *Applied deep learning: A case-based approach to understanding deep neural networks*. New York, NY: Apress.

Mizera, I., Müller, C. H., Hubert, M., Rousseeuw, P. J., Branden, K. V., Serfling, R., ... McCullagh, P. (2004). Location- scale depth [with comments, rejoinder]. *Journal of the American Statistical Association*, *99*(468), 949–989.

Morris, G. L., & Fan, J. (1998). *Reservoir sedimentation handbook* (Vol. 1.01). Pennsylvania, PA: McGraw-Hill Book Co.

Nisbet, R., Miner, G., & Yale, K. (2017). *Handbook of statistical analysis and data mining applications* (2nd ed.). Orlando, FL: Academic Press, Inc.

Okano, M., Takanagi, J., Fujii, T., & Ando, A. (2004). Study of reservoir sedimentation management based on actual sediment inflow: Thecase of de- velopment of the sediment yield potential map. *Journal of Japan Society of Dam Engineers*, *14*(3), 167–176.

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*. Berlin and Heidelberg: Springer.

Park, S., & Kwak, N. (2017). Analysis on the dropout effect in convolutional neural networks. In *Computer Vision – ACCV 2016: 13th Asian conference on computer vision, Taipei, Taiwan, China* (pp. 189–204).

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, E. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Perez, C. E. (2017). Is deep learning innovation just due to brute force? Retrieved from https://medium.com/intuitionmachine/the-brute-force-method-of-deep-learning-innovation-58b497323ae5

Pontius, R., Thontteh, O., & Chen, H. (2008). Components of information for multiple resolution comparison between maps that share a real variable. *Environmental and Ecological Statistics*, *15*, 111–142.

Rakhmatullaev, S. (2010). *Improvement of operational methods for the assessment of the water reservoir useful storage capacity using geoinformation systems*. Case study of the Akdarya Reservoir, Samarqand Province, Uzbekistan. Université Sciences et Technologies - Bordeaux I.

Reza Rahmanian, M., & Banihashemi, M. (2012). Characterization of sedimentation pattern in reservoirs. *Canadian Journal of Civil Engineering*, *39*, 951–956.

Schleiss, A., De Cesare, G., & Althaus, J. J. (2010). Verlandung der Stauseen gefährdet die nachhaltige Nutzung der Wasserkraft. *Luft, Wasser, Energie.*, *6*, 595–614.

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.

Shah, D., Campbell, W., & Zulkernine, F. H. (2018). A comparative study of lstm and dnn for stock market forecasting. *2018 IEEE International Conference on Big Data (Big Data)* (pp. 4148–4155). Washington, DC.

Simoes, F. J. M., & Yang, C. T. (2006). *Sedimentation modeling for rivers and reservoirs*. Reston, VA: USGS.

Spiess, A.-N., & Neumeyer, N. (2010). An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach. *BMC Pharmacology, 10*(6). doi:10.1186/1471-2210-10-6

Sumi, T. (2015). *Comprehensive reservoir sedimentation countermeasures in Japan*. International Workshop on Sediment Bypass Tunnels, Zurich.

Um, T. T., Pfister, F. M., Pichler, D., Endo, S., Lang, M., Hirche, S., . . . Kulic, D. (2017). Data augmentation of wearable sensor data for Parkinson's disease monitoring using convolutional neural networks. *Proceedings of 19th ACM International Conference on Multimodal Interaction (ICMI'17)*, Glasgow: ACM.

Velasco, J. M., Garnica, O., Lanchares, J., Botella, M., & Ignacio Hidalgo, J. (2018). *Combining data augmentation, EDAs and grammatical evolution for blood glucose forecasting*. Berlin and Heidelberg: Springer.

World Commission on Environment and Development. (2018). *Report of the world commission on environment and development: Our common future*. Oxford: Oxford University Press.

Xiao, G., & Xu, X. (2012). Study on bezier curve variable step-length algo- rithm. *Physics Procedia*, *25*, 1781–1786. International Conference on Solid State Devices and Materials Science, April 1–2, 2012, Macao.

Yang, S.-Q. (2013). Why cannot sediment transport be accurately predicted. *35th IAHR World Congress*. Chengdu, China: IAHR.

Zarfl, C., Lumsdon, A. E., Berlekamp, J., Tydecks, L., & Tockner, K. (2015). A global boom in hydro-power dam construction. *Aquatic Sciences*, *77*, 1015–1621.

Zeleke, 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: Science, Policy and Management for Sustainable Use, 18*(4), 366–371. doi:10.1111/lre.12047.

Zhang, C., Vinyals, O., Munos, R., & Bengio, S. (2018). A study on overfitting in deep reinforcement learning. *CoRR, abs/1804.06893*. Retrieved from http://arxiv.org/abs/1804.06893